

B7

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number  
WO 02/20836 A2

- (51) International Patent Classification<sup>7</sup>: C12Q 1/68 (74) Agents: GARDNER, Rebecca et al.; Frank B. Dehn & Co., 179 Queen Victoria Street, London EC4V 4EL (GB).
- (21) International Application Number: PCT/GB01/04015
- (22) International Filing Date:  
7 September 2001 (07.09.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
0021977.4 7 September 2000 (07.09.2000) GB
- (71) Applicant (for all designated States except US): PY-ROSEQUENCING AB [SE/SE]; Vallongatan 1, S-752 28 Uppsala (SE).
- (71) Applicant (for GB only): GARDNER, Rebecca [GB/GB]; Frank B. Dehn & Co., 179 Queen Victoria Street, London EC4V 4EL (GB).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): NYREN, Pål [SE/SE]; c/o Department of Biotechnology, KTH, SCFAB, S-106 91 Stockholm (SE). RONAGHI, Mostafa [SE/US]; Stanford DNA Sequencing & Technology Center, 855 California Avenue, Palo Alto, CA 94304 (US). TALLSJÖ, Annika [SE/SE]; PyroSequencing AB, Vallongatan 1, S-752 28 Uppsala (SE).
- (81) Designated States (national): AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EC, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/20836 A2

(54) Title: METHOD OF SEQUENCING DNA

(57) Abstract: The present invention provides a method of identifying a base at a target position in a sample nucleic acid sequence, said method comprising: subjecting a primer hybridised to said sample nucleic acid immediately adjacent to the target position, to a polymerase primer extension reaction in the presence of a nucleotide, whereby the nucleotide will only become incorporated if it is complementary to the base in the target position, and determining whether or not said nucleotide is incorporated by detecting whether Ppi is released, the identity of the target base being determined from the identity of any nucleotide incorporated, wherein, where said nucleotide comprises an adenine base, an  $\alpha$ -thio triphosphate analogue of said nucleotide is used, and the Rp isomer of said analogue and/or the degradation products of said analogue are eliminated from the polymerase reaction step.

### Method of sequencing DNA

This invention relates to improvements in methods of sequencing DNA, based on the detection of base incorporation by the release of pyrophosphate (PPi).

DNA sequencing is an essential tool in molecular genetic analysis. The ability to determine DNA nucleotide sequences has become increasingly important as efforts have proceeded to determine the sequences of the large genomes of humans and other organisms.

Techniques enabling the rapid detection of a single DNA base change, or a few base changes, are also important tools for genetic analysis, for example in clinical situations in the analysis of genetic diseases or certain cancers. Indeed, as more and more diseases are discovered to be associated with changes at the genetic level, most notably single nucleotide polymorphisms (SNPs), the need for methods of both screening for SNPs or other mutations or genetic changes (by sequencing representative genomic samples) and scoring SNPs (or other mutations/changes) grows. Thus, as well as the development of novel sequencing technologies for determining the sequence of longer stretches of DNA, the art has also seen a rapid rise in the development of technologies for detecting single (or a few) base changes. Such protocols to determine more limited sequence information, relating to only one or a few bases are termed mini-sequencing.

The method most commonly used as the basis for DNA sequencing, or for identifying a target DNA base, is the enzymatic chain-termination method of Sanger. Traditionally, such methods relied on gel electrophoresis to resolve, according to their size, DNA fragments produced from a larger DNA segment. However, in recent years various sequencing technologies have evolved which rely on a range of different detection

- 2 -

strategies, such as mass spectrometry and array technologies.

One class of sequencing methods assuming importance in the art are those which rely upon the detection of PPi release as the detection strategy. It has been found that such methods lend themselves admirably to large scale genomic projects or clinical sequencing or screening, where relatively cost-effective units with high throughput are needed.

Methods of sequencing based on the concept of detecting inorganic pyrophosphate (PPi) which is released during a polymerase reaction have been described in the literature for example (WO 93/23564, WO 89/09283, WO98/13523 and WO 98/28440). As each nucleotide is added to a growing nucleic acid strand during a polymerase reaction, a pyrophosphate molecule is released. It has been found that pyrophosphate released under these conditions can readily be detected, for example enzymically e.g. by the generation of light in the luciferase-luciferin reaction. Such methods enable a base to be identified in a target position and DNA to be sequenced simply and rapidly whilst avoiding the need for electrophoresis and the use of labels.

At its most basic, a PPi-based sequencing reaction involves simply carrying out a primer-directed polymerase extension reaction, and detecting whether or not that nucleotide has been incorporated by detecting whether or not PPi has been released. Conveniently, this detection of PPi-release may be achieved enzymatically, and most conveniently by means of a luciferase-based light detection reaction termed ELIDA (see further below).

It has been found that dATP added as a nucleotide for incorporation, interferes with the luciferase reaction used for PPi detection. Accordingly, a major improvement to the basic PPi-based sequencing method has been to use, in place of dATP, a dATP analogue

- 3 -

(specifically dATP $\alpha$ s) which is incapable of acting as a substrate for luciferase, but which is nonetheless capable of being incorporated into a nucleotide chain by a polymerase enzyme (WO98/13523).

Further improvements to the basic PPI-based sequencing technique include the use of a nucleotide degrading enzyme such as apyrase during the polymerase step, so that unincorporated nucleotides are degraded, as described in WO 98/28440, and the use of a single-stranded nucleic acid binding protein in the reaction mixture after annealing of the primers to the template, which has been found to have a beneficial effect in reducing the number of false signals, as described in WO00/43540.

However, even with the modified and improved PPI-based sequencing methods mentioned above, there is still room for improvement, for example to increase the efficiency and/or accuracy of the procedure, or, as discussed further below, to increase the sequence read length possible. The present invention addresses these needs.

In particular, the present invention is concerned with methods of PPI-based sequencing which use an  $\alpha$ -thio analogue of deoxy ATP (dATP) (or dideoxy ATP (ddATP)) namely an (1-thio) triphosphate (or  $\alpha$ -thiophosphate) analogue of deoxy or dideoxy ATP, preferably deoxyadenosine [1-thio] triphosphate or deoxyadenosine  $\alpha$ -thiotriphosphate (dATP $\alpha$ s) as it is also known. dATP $\alpha$ s (as with all  $\alpha$ -thio nucleotide analogues) occurs as a mixture of isomers, the Rp isomer and the Sp isomer.

When dATP $\alpha$ S (and/or other  $\alpha$ -thio nucleotides) are used, it has been found that the efficiency of the sequencing method decreases as the number of cycles increases, and in particular that the read length attainable is limited (e.g. to 40-50 bases). This is believed to be due to the accumulation of inhibitory substances in the reaction system. The present

- 4 -

invention is particularly concerned with reducing or removing such inhibitory effects.

More particularly, it has surprisingly been found that removing or excluding the Rp isomer of  $\alpha$ -thio nucleotide analogues and/or the degradation products of said analogue from the polymerisation mixture improves the efficiency of PPI-based sequencing methods, and in particular, such methods based on detecting PPI by a luciferase-based ELIDA reaction.

Whilst not wishing to be bound by theory, it is believed that a number of reasons or effects may contribute to this beneficial effect of Rp isomer and/or the degradation products of the  $\alpha$ -thio nucleotide analogue elimination on the sequencing reaction. Certain of those effects are explained further below, but in particular we believe that the Rp isomer and/or the degradation products of the  $\alpha$ -thio nucleotide analogue are capable of inhibiting polymerase activity. Thus, it is believed that one such contributory effect may be due to the effect of the Rp isomer and/or the degradation products of the  $\alpha$ -thio nucleotide analogue, not hitherto appreciated, in inhibiting the activity of polymerase. Hence, by excluding or removing the Rp isomer and/or the degradation products of the  $\alpha$ -thio nucleotide analogue, this inhibitory effect may be avoided, leading to a more efficient and faster polymerisation reaction, and more even signals generated in the PPI-detection reaction. Further, by removing or excluding the Rp isomer and/or the degradation products of the  $\alpha$ -thio nucleotide analogue, the fidelity of the DNA synthesis is increased, i.e. the number of misincorporation events is decreased.

In one aspect, the present invention thus provides a method of identifying a base at a target position in a sample nucleic acid sequence, said method comprising:

subjecting a primer hybridised to said sample nucleic acid immediately adjacent to the target

- 5 -

position, to a polymerase primer extension reaction in the presence of a nucleotide, whereby the nucleotide will only become incorporated if it is complementary to the base in the target position, and determining whether or not said nucleotide is incorporated by detecting whether PPi is released, the identity of the target base being determined from the identity of any nucleotide incorporated,

wherein, where said nucleotide comprises an adenine base, an  $\alpha$ -thio triphosphate analogue of said nucleotide is used, and the Rp isomer of said analogue and/or the degradation products of said analogue are eliminated from the polymerase reaction step.

Use of  $\alpha$ -trio triphosphate analogue of a nucleotide comprising an adenine base and eliminating the Rp isomer thereof and/or eliminating the degradation products of said analogue are essential features of the present invention, which this does not relate to, e.g. minisequencing reactions where only nucleotides comprising guanine, thymine or cytosine bases are added for incorporation.

Subject to the proviso that where the nucleotide comprises an adenine base (A), an  $\alpha$ -thio triphosphate analogue of said nucleotide is used, the nucleotide may be any nucleotide capable of incorporation by a polymerase enzyme into a nucleic acid chain or molecule. Thus, for example, the nucleotide may be a deoxynucleotide (dNTP, deoxynucleoside triphosphate) or dideoxynucleotide (ddNTP, dideoxynucleoside triphosphate).

Conveniently, for sequencing purposes, guanine (G), cytosine (C), thymine (T) or adenine (A) deoxy- or dideoxy-nucleotides may be used. Thus the nucleotide may be dGTP (deoxyguanosine triphosphate), dCTP (deoxycytidine triphosphate) or dTTP (deoxythymidine triphosphate) or in place of dATP (deoxyadenosine triphosphate) its  $\alpha$ -thiotriphosphate analogue, dATP $\alpha$ s

- 6 -

(deoxyadenosine  $\alpha$ -thiotriphosphate or deoxyadenosine[1-thio]-triphosphate as it is also known). Analogously, the nucleotides may be ddGTP, ddCTP, or dTTP, or in place of ddATP, ddATP $\alpha$ s.

The present invention thus requires the use of an  $\alpha$ -thiotriphosphate analogue of an adenine nucleotide, but for the other bases (G, T or C), a native nucleotide may be used (or indeed any other nucleotide, e.g. nucleotide derivative, it is desired to use, provided that it can be incorporated by a polymerase enzyme). Thus, according to the present invention, at least an  $\alpha$ -thio analogue of an adenine nucleotide is used.

However, it may in certain cases be desirable also to use  $\alpha$ -thio analogues of one or more other nucleotides (i.e. of guanine, cytosine or thymine nucleotides). In any such case where an  $\alpha$ -thio analogue of a nucleotide is used, then according to the present invention the Rp isomer of said analogue is eliminated from the polymerase reaction step. In other words, in the method of the invention, where the nucleotide is an  $\alpha$ -thio nucleotide (e.g. deoxy- or dideoxy-nucleoside  $\alpha$ -thiotriphosphate (dNTP $\alpha$ S or ddNTP $\alpha$ S)), the Rp isomer of said  $\alpha$ -thio nucleotide is eliminated from the polymerase reaction step, and /or the degradation products of the NTP $\alpha$ S are eliminated.

For convenience, the term "deoxynucleoside  $\alpha$ -thiotriphosphate" (dNTP $\alpha$ S) as used herein thus includes deoxyadenosine  $\alpha$ -thiotriphosphate (dATP $\alpha$ S), deoxycytidine  $\alpha$ -thiotriphosphate (dCTP $\alpha$ S), deoxyguanosine  $\alpha$ -thiotriphosphate (dGTP $\alpha$ ) and deoxythymidine  $\alpha$ -thiotriphosphate (dTTP $\alpha$ S). Analogously, the term "dideoxy nucleotide  $\alpha$ -thio triphosphate" includes the dideoxy equivalent. The term "dideoxynucleotide" as used herein includes all 2'-deoxynucleotides in which the 3'-hydroxyl-group is absent or modified and thus, whilst being able to be added to the primer in the presence of the polymerase,

- 7 -

is unable to enter into a subsequent polymerisation reaction, i.e. a dideoxy nucleotide is thus a "chain terminator", and as is well known in the art, certain sequencing methods may employ such chain-terminating nucleotides.

The term "NTP $\alpha$ S" (nucleoside  $\alpha$ -thiotriphosphate) is used herein to refer to all  $\alpha$ -thiotriphosphate nucleotide analogues (i.e.  $\alpha$ -thionucleotides) which may be used according to the present invention, and includes both ribo- and deoxyribo- (or dideoxy-) nucleotides. These include, most notably, dNTP $\alpha$ S and ddNTP $\alpha$ S.

When synthesized,  $\alpha$ -thiotriphosphate nucleotide analogues (e.g. dd- or dNTP $\alpha$ S) are produced typically in two isomeric forms, the Sp and Rp isomers. An  $\alpha$ -thiotriphosphate nucleotide analogue possesses a chiral centre, and thus the 2 species are enantiomers. The Sp isomer is the left-handed isomer, also designated the L isomer. The right-handed isomer is the Rp isomer, also known as the D-isomer (see e.g. Eckstein (1985), Ann. Rev. Biochem., 54: (367-402)).

The Rp isomer of NTP $\alpha$ S does not act as a substrate for a polymerase enzyme, as used in the primer extension reaction. Surprisingly, however, as mentioned above, it has been found that use of the Rp isomer of NTP $\alpha$ S leads to inhibition of enzymes involved in the PPi-based sequencing reactions. The precise nature of these various inhibitory effects is not entirely clear, but it is believed that the Rp isomer of NTP $\alpha$ S itself (and possibly also the Rp isomer of NDP $\alpha$ S and Rp isomer of NMP $\alpha$ S) are responsible for the inhibitory effects observed, including, as mentioned above, inhibition of polymerase. The degradation products of NTP $\alpha$ S are also responsible for the inhibitory effects seen. The degradation products include, but are not limited to, NDP $\alpha$ S, NMP $\alpha$ S and any degradation products thereof. The inhibitory effects of the Rp isomer of dATP $\alpha$ S on nucleotide degrading enzymes and PPi detection enzymes



- 8 -

have been investigated (see Example 2) and figures 8I, 8J, 8K and 8L) and it has been shown that the Rp isomer and/or degradation products of the dATP $\alpha$ S do inhibit enzymes involved in nucleotide degradation and detection of PPi release.

It will be appreciated that when the target base immediately 3'- of the primer has an identical base 3'- thereto, and the polymerisation is effected with a deoxynucleotide (rather than a dideoxynucleotide) the extension reaction will add two bases at the same time and indeed any sequence of successive identical bases in the sample will lead to simultaneous incorporation of corresponding bases into the primer. However, the amount of pyrophosphate liberated will clearly be proportional to the number of incorporated bases so that there is no difficulty in detecting such repetitions.

Since the primer is extended by a single base by the procedure described above (or by a sequence of identical bases), the extended primer can serve in exactly the same way in a repeated procedure to determine the next base in the sequence, thus permitting the whole sample to be sequenced.

The method of the invention may thus be used to determine the identity (i.e. sequence) of a single base. However, conveniently, by repeating the primer extension steps in the presence of a further (successive) nucleotide, the sequence (or identity) of a further base in the sequence of the sample nucleic acid may be revealed. Accordingly, the method of the invention may be used to determine the identity of one or more bases in a sample nucleic acid (i.e. to determine the sequence of one or more bases in a sample nucleic acid).

The method of the invention thus has utility in a number of different sequencing methods and formats, including mini-sequencing procedures e.g. detection of single base changes (for example, in detecting point mutations, or polymorphisms, or allelic variations etc).

- 9 -

Accordingly, the method of the invention may thus be used in a "full" sequencing procedure, i.e. the identification of the sequential order of the bases in a stretch of nucleotides, as well in single base detection procedures.

For example, to determine sequence information in a target nucleotide sequence (i.e. target or sample nucleic acid), different nucleotides may be added either to separate aliquots of sample-primer mixture (e.g. four aliquots, one for each of the four, A, T, G or C nucleotides) or successively to the same sample-primer mixture and subjected to the polymerase reaction to indicate which nucleotide is incorporated.

In order to sequence the sample nucleic acid, the procedure may be repeated one or more times i.e. cyclically, as is known in the art. In this way the identity of several or many bases in the sample nucleic acid may be identified, essentially in the same reaction.

Where separate aliquots are used, once it has been identified what base has been incorporated (i.e. in which aliquot incorporation has taken place), the "incorporated" base may be added to the "unreacted" aliquots, to extend the primer in all aliquots, before repeating the process (cycling) to sequence the next base. In the "successive" embodiment, a different nucleotide may be added successively until incorporation is indicated by PPi release, whereupon the procedure may be repeated.

Hence, a sequencing protocol may involve annealing a primer as described above, adding a nucleotide, performing a polymerase-catalysed primer extension reaction, detecting the presence or absence of incorporation of said nucleotide by detecting any PPi released, and repeating the nucleotide addition and primer extension steps etc. one or more times. As discussed above, single (i.e. individual) nucleotides

- 10 -

may be added successively to the same primer-template mixture, or to separate aliquots of primer-template mixture, etc. according to choice, and the sequence information it is desired to obtain.

In order to permit the repeated or successive (iterative) addition of nucleotides in a multiple-base sequencing procedure, the previously-added nucleotide must be removed. This may be achieved by washing, or more conveniently, by using a nucleotide-degrading enzyme, for example as described in detail in WO98/28440.

Accordingly, in a principal embodiment of the present invention, a nucleotide degrading enzyme is used to degrade any unincorporated or excess nucleotide. Thus, if a nucleotide is added which is not incorporated (because it is not complementary to the target base), or any added nucleotide remains after an incorporation event (i.e. excess nucleotides) then such unincorporated nucleotides may readily be removed by using a nucleotide-degrading enzyme. This is described in detail in WO98/28440.

As will be discussed in more detail below, it has been observed that inhibitory effects due to the use of an  $\alpha$ -thio nucleotide particularly occur (or are observed) when a nucleotide degrading enzyme is used, and that such effects may be beneficially abrogated according to the present invention, by the methods described herein.

The term "nucleotide degrading enzyme" as used herein includes any enzyme capable of specifically or non-specifically degrading nucleotides, including at least nucleoside triphosphates (NTPs), but optionally also di- and mono-phosphates, and any mixture or combination of such enzymes, provided that a nucleoside triphosphatase or other NTP-degrading activity is present. Although nucleotide-degrading enzymes having a phosphatase activity may conveniently be used according

- 11 -

to the invention, any enzyme having any nucleotide or nucleoside degrading activity may be used, e.g. enzymes which cleave nucleotides at positions other than at the phosphate group, for example at the base or sugar residues. Thus, a nucleoside triphosphate degrading enzyme is essential for the invention. Nucleoside di- and/or mono-phosphate degrading enzymes are optional and may be used in combination with a nucleoside triphosphate degrading enzyme. A phosphatase used as a nucleotide degrading enzyme according to this aspect of the invention should meet several criteria, notably the inhibitory constant ( $K_i$ ) for the phosphate ( $P_i$ ) product of phosphatase action, should not be too low, so that the enzyme is not inhibited by accumulating phosphate. Secondly, the enzyme needs to act relatively fast and should not be too slow (certain phosphatase enzymes can act too slowly to be practical) and thirdly it should degrade all four nucleotide substrates (i.e. A, T, G and C substrates) with more or less equal efficiency. As discussed further below, ATP generated in the ELIDA reactions is preferred for  $PP_i$  detection, and a nucleotide degrading enzyme useful in the invention should also efficiently degrade ATP. This leads to an efficient "turning-off" of the signal. It will be noted that not all phosphatase enzymes (e.g. alkaline phosphatases that are strongly inhibited by the product phosphate) meet these criteria, and so not all may be suitable for use as a nucleotide degrading enzyme according to the invention. However, such suitability may readily be assessed by routine experiments. The preferred nucleotide degrading enzyme is apyrase, which is both a nucleoside diphosphatase and triphosphatase, catalysing the reactions  $NTP \rightarrow NDP + P_i$  and  $NDP \rightarrow NMP + P_i$  (where NTP is a nucleoside triphosphate, NDP is a nucleoside diphosphate, NMP is a nucleotide monophosphate and  $P_i$  is inorganic phosphate). Apyrase may be obtained from the Sigma Chemical Company. Other

- 12 -

possible nucleotide degrading enzymes include Pig Pancreas nucleoside triphosphate diphosphorydrolase (Le Bel et al., 1980, J. Biol. Chem., 255, 1227-1233).

Further enzymes are described in the literature.

The nucleotide-degrading enzyme may conveniently be included during the polymerase (i.e. primer extension) reaction step. Thus, for example the polymerase reaction may conveniently be performed in the presence of a nucleotide-degrading enzyme. Although less preferred such an enzyme may also be added after nucleotide incorporation (or non-incorporation) has taken place, i.e. after the polymerase reaction step.

Thus, the nucleotide-degrading enzyme (e.g. apyrase) may be added to the polymerase reaction mixture (i.e. sample nucleic acid, primer and polymerase) in any convenient way, for example prior to or simultaneously with initiation of the reaction, or after the polymerase reaction has taken place, e.g. prior to adding nucleotides to the sample/primer/polymerase to initiate the reaction, or after the polymerase and nucleotide are added to the sample/primer mixture.

Conveniently, the nucleotide-degrading enzyme may simply be included in the reaction mixture for the polymerase reaction, which may be initiated by the addition of the nucleotide.

A further surprising feature of the present invention is the observation that use of the Rp isomer of  $\alpha$ -thio nucleotides and/or degradation products of said  $\alpha$ -thio nucleotides lead to inhibition of a nucleotide-degrading enzyme of the present invention, most notably apyrase. It is believed that the Rp isomer is not capable of acting as a substrate for apyrase, and accordingly that it is not degraded by the nucleotide-degrading activity of the apyrase. Alternatively, products of NTP $\alpha$ S degradation may be inhibitory. Thus, a situation may be created of successive accumulation of the inhibitory and inactive Rp isomer (or of inhibitory

- 13 -

degradation products) during the sequencing procedure. Thus, a further benefit derivable from the elimination of the Rp isomer and/or degradation products of said  $\alpha$ -thio nucleotides according to the present invention, is that inhibition of the apyrase may be reduced or avoided. This has very significant, and heretofore unpredictable, benefits on the efficiency and performance of the sequencing method, and represents a significant and surprising advantage of the present invention.

Degradation of the unincorporated nucleotides by apyrase has the benefit of producing even and well defined (e.g. narrow/defined or sharp) PPi detection signals (see further below). Where such degradation is inefficient, due to inhibition of the apyrase enzyme, this benefit is progressively lost. Thus, inhibition of apyrase is seen, directly or indirectly, as a slower nucleotide degradation rate, and consequently as a "wider" or less well-defined PPi-detection signal, in the later cycles of sequencing. In particular, in the context of ELIDA detection of PPi release, described further below, the decay in signal observed (light) is "seen" as degradation of ATP (generated in the ELIDA reactions). Thus, a slower rate of degradation of ATP is "seen". It can be inferred from this that the rate of unincorporated nucleotide degradation is also slower. Further, non-synchronised extension may result. This occurs, since the presence of undegraded unincorporated nucleotides may lead to multiple extension reactions occurring out of phase, and leading to overlapping signals (signal overlay or out of phase signals). Such non-synchronised extension thus limits the number of nucleotides which may be sequenced in a given sequencing run i.e. the read-length attainable.

The ability to sequence longer stretches of nucleic acid is a desirable goal in the sequencing field. The inhibitory effects resulting from the use of  $\alpha$ -thio

- 14 -

nucleotides on the enzymes present within the PPi-based sequencing methods of the present invention, leads to decreased fidelity and non-synchronised extension, thus limiting the read length attainable, i.e. the length of the nucleic acid which can be successfully sequenced. Read-length may be improved according to the present invention by eliminating the Rp isomer of the  $\alpha$ -thionucleotide and/or degradation product of  $\alpha$ -thio nucleotides.

Attainable read-length is to some extent a parameter dependent upon the acceptance criteria adopted. Cleanness of signal may be lost, but the sequence may nonetheless still be readable to a skilled and experienced practitioner. Precise read-length limits therefore are not always meaningful or cannot be applied generally, and may depend on circumstances. However, it has in certain cases been found that 50 or more, or even 100 or more (e.g. 200 or more) bases may readily be read according to the present invention. The methods of the invention are thus suitable for the sequencing of 100 bases or more. In particular, the present invention may advantageously be used in the sequencing of 50 or more, 60 or more, 70 or more or 80 or more bases.

PPi release may be detected according to the present invention in any desired or convenient way. PPi can be determined by many different methods and a number of enzymatic methods have been described in the literature (Reeves et al., (1969), Anal. Biochem., 28, 282-287; Guillory et al., (1971), Anal. Biochem., 39, 170-180; Johnson et al., (1968), Anal. Biochem., 15, 273; Cook et al., (1978), Anal. Biochem. 91, 557-565; and Drake et al., (1979), Anal. Biochem. 94, 117-120).

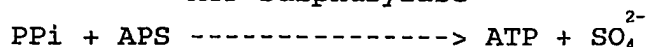
It is preferred to use a luciferase-based (e.g. a luciferin luciferase-based) light generating reaction to detect the release of pyrophosphate since the amount of light generated is substantially proportional to the

- 15 -

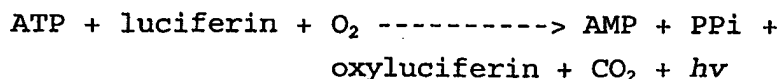
amount of pyrophosphate released which, in turn, is directly proportional to the number of bases incorporated. The amount of light can readily be estimated by a suitable light sensitive device such as a luminometer.

Luciferase-based reactions to detect the release of PPi are well known in the art. In particular, a method for detecting PPi release based on the enzymes ATP sulphurylase and luciferase has been developed by Nyrén and Lundin (Anal. Biochem., 151, 504-509, 1985) and termed ELIDA (Enzymatic Luminometric Inorganic Pyrophosphate Detection Assay). The use of the ELIDA method to detect PPi is preferred according to the present invention. The method may however be modified, for example by the use of a more thermostable luciferase (Kaliyama *et al.*, 1994, Biosci. Biotech. Biochem., 58, 1170-1171) and/or ATP sulfurylase (Onda *et al.*, 1996, Bioscience, Biotechnology and Biochemistry, 60:10, 1740-42). This method is based on the following reactions:

ATP sulphurylase



luciferase



(APS = adenosine 5'-phosphosulphate)

Accordingly, it is preferred to detect PPi release enzymatically, and the preferred detection enzymes used in the PPi detection reaction are ATP sulphurylase and luciferase.

In a PPi detection reaction based on the enzymes ATP sulphurylase and luciferase, the signal (corresponding to PPi released) is seen as light. The generation of the light can be observed as a curve known



- 16 -

as a pyrogram. Light is generated by luciferase action on the product, ATP (produced by a reaction between PPI and APS (see below) mediated by ATP sulphurylase) and, where a nucleotide-degrading enzyme such as apyrase is used, this light generation is then "turned off" by the action of the nucleotide-degrading enzyme, degrading the ATP which is the substrate for luciferase. The slope of the ascending curve may be seen as indicative of the activities of DNA polymerase (PPI release) and ATP sulphurylase (generating ATP from the PPI, thereby providing a substrate for luciferase). The height of the signal is dependent on the activity of luciferase, and the slope of the descending curve is, as explained above, indicative of the activity of the nucleotide-degrading enzyme.

Advantageously, by including the PPI detection enzyme(s) (i.e. the enzyme or enzymes necessary to achieve PPI detection according to the enzymatic detection system selected, which in the case of ELIDA, will be ATP sulphurylase and luciferase) in the polymerase reaction step, the method of the invention may readily be adapted to permit the sequencing reactions to be continuously monitored in real-time, with a signal being generated and detected, as each nucleotide is incorporated. The benefits of such an approach are discussed in more detail in WO98/13523.

Thus, the PPI detection enzymes (along with any enzyme substrates or other reagents necessary for the PPI detection reaction) may simply be included in the polymerase reaction mixture.

More particularly, to carry out this embodiment of the method of the invention, the detection enzymes are included in the polymerase reaction step i.e. in the chain extension reaction step. Thus the detection enzymes are added to the reaction mix for the polymerase step prior to, simultaneously with or during the polymerase reaction. In the case of an ELIDA detection

- 17 -

reaction, the reaction mix for the polymerase reaction may thus include at least one nucleotide, polymerase, luciferin, APS, ATP sulphurylase and luciferase. The polymerase reaction may be initiated by addition of the polymerase or, more preferably the nucleotide, and preferably the detection enzymes are already present at the time the reaction is initiated, or they may be added with the reagent that initiates the reaction. If a nucleotide degrading enzyme such as apyrase is used, it will of course be understood that polymerase addition should not follow apyrase, if nucleotides are present.

The present invention thus permits PPi release to be detected during the polymerase reaction giving a real-time signal. The sequencing reactions may be continuously monitored in real-time.

The Rp isomer and/or the degradation products of NTP $\alpha$ S may be eliminated according to the present invention in a number of ways, and the term "eliminated" requires simply that the Rp isomer and/or degradation products be removed or excluded from the polymerase reaction step. This may be achieved in any convenient way. For example, the Rp isomer may be excluded, by not being present in the NTP $\alpha$ S preparation which is added to or included in the polymerase reaction mixture. Alternatively, the NTP $\alpha$ S preparation may initially include the Rp isomer, but it may be removed, before, during or after, inclusion in or addition to the polymerase reaction mixture, for example by enzymic degradation. Combinations of means for eliminating the Rp isomer (e.g. removing and excluding) may also be used. Thus, the Rp isomer may be removed or excluded from the polymerase reaction step, from the step after polymerisation has occurred (i.e. the nucleotide degradation step) or the detection of nucleotide incorporation step.

The degradation products of the NTP $\alpha$ -S may be removed during the polymerisation step, the nucleotide

- 18 -

degrading step or the detection of nucleotide incorporation step. For example the degradation products are eliminated by enzymic degradation.

In one embodiment of the invention, the Rp isomer may thus be eliminated by using a preparation of NTP $\alpha$ S which contains only the Sp isomer i.e. by using the Sp isomer of NTP $\alpha$ S only e.g. pure Sp isomer. Individual Rp and Sp isomers of dATP $\alpha$ S are available commercially (e.g. from Biolog Life Science, Bremen, DE).

Alternatively, such isomers may readily be synthesised stereo-specifically, or purified (separated) from a racemic mixture using techniques well known in the art and described in the literature (see e.g. Eckstein, *supra*). dATP $\alpha$ S, along with the  $\alpha$ -thio analogues of dCTP, dGTP and dTTP, may be purchased from Amersham Pharmacia Biotech (Uppsala, SE).

The Rp isomer may be reduced in the reaction step to less than 10% of the total dATP $\alpha$ S, preferably less than 5%, more preferably less than 3% Rp isomer.

In an alternative embodiment, the Rp isomer and/or degradation products of the  $\alpha$ -thio nucleotide analogue may be removed by using an enzyme which degrades it, notably alkaline phosphatase. In particular, preliminary experimental results lead us to believe that alkaline phosphatase is capable of degrading both isomers of NTP $\alpha$ S, including the Rp isomer and degradation products of NTP $\alpha$ S. Thus, Rp isomer and degradation products of NTP $\alpha$ S present in the reaction mixture, or in the NTP $\alpha$ S preparation, may readily be removed (degraded) by adding an alkaline phosphatase enzyme.

In a more particular embodiment, a combination of Sp isomer together with an alkaline phosphatase enzyme may be used.

Conveniently, the alkaline phosphatase may simply be included during the polymerase reaction step. This may be achieved by adding the enzyme to the polymerase

- 19 -

reaction mixture prior to, or simultaneously with, initiation of the polymerase reaction, or after the polymerase reaction has taken place, e.g. prior to adding nucleotides to the sample/primer/polymerase mixture to initiate the reaction, or after the polymerase and nucleotide are added to the sample/primer mixture.

Alternatively, the alkaline phosphatase can be immobilized on a solid support e.g. a particulate solid support (e.g. magnetic beads) or a filter, or dipstick etc. and it may be added to the polymerase reaction mixture at a convenient time. When the Rp isomer has been degraded, the immobilised enzyme may be removed from the reaction mixture (e.g. it may be withdrawn or captured, e.g. magnetically in the case of magnetic beads), before the next nucleotide is added. The procedure may then be repeated to sequence more bases.

Alkaline phosphatase catalyses the removal of 5'-phosphate residues from nucleoside tri-, di- and mono-phosphates (including ribo-NTPs and dNTPs and, as mentioned above the both isomers of NTPs). Thus, not only are nucleotides (nucleoside triphosphates) degraded but also NDPs and NMPs. Shrimp, bacterial and/or calf-intestinal alkaline phosphatase are the preferred enzymes, but any suitable alkaline phosphatase enzyme may be used in the method of the invention, from any convenient source. The preferred enzyme is shrimp alkaline phosphatase. There are many commercial sources of alkaline phosphatase enzymes, or they may, if desired, be isolated from a producing organism.

The amount of alkaline phosphatase enzyme to be used will depend upon the precise reaction system used, reaction conditions etc. and can readily be determined by routine experiments. It has been found, for example, that concentrations of 50 mU to 10 U may be employed.

As mentioned above, alkaline phosphatase has an activity in degrading NDPs and NMPs as well as NTPs. It

- 20 -

further has the benefit of acting on  $\alpha$ -thio modified substrates also, thus including not only NTP $\alpha$ S, but also NDP $\alpha$ S and NMP $\alpha$ S. This activity lends a further benefit to the present invention. Thus, it has been observed that the enzyme apyrase, and also the enzyme luciferase used in the PPi detection method preferred according to the present invention, are inhibited when an NTP $\alpha$ S is used for nucleotide incorporation. The precise nature of the inhibitory effects observed is unclear, but may be due to NTP $\alpha$ S product or substrate inhibition. Thus, NTP $\alpha$ S may itself be inhibitory, or more likely a degradation product of NTP $\alpha$ S (for example resulting from apyrase action) is inhibitory. For example, it is postulated that the degradation products NDP $\alpha$ S and NMP $\alpha$ S, (in particular dADP $\alpha$ S and dAMP $\alpha$ S) produced by the action of apyrase on NTP $\alpha$ S added to the polymerase reaction mixture for incorporation, may be inhibitory. More particularly, it is believed that both the Sp and Rp isomers of NDP $\alpha$ S and/or NMP $\alpha$ S inhibit luciferase, and also the enzyme apyrase (see further in the Examples below). Such inhibitory substances may be removed (or reduced) by the action of alkaline phosphatase.

As mentioned above, apyrase inhibition is indicated by a slower nucleotide degradation rate, and hence an increased signal width. Non-synchronised extension may also occur, and this limits the read length attainable. Luciferase inhibition is indicated by a decrease in signal intensity. Thus, as shown further in the Examples below, inhibition of luciferase can be seen as a steady decrease in signal peak height as the nucleotides are incorporated in the later cycles of sequencing.

The effect of decreased signal intensity (luciferase inhibition) may combine with the effect of loss of signal definition (e.g. increased signal width or out of phase signals) (apyrase inhibition) to reduce the efficiency of the sequencing method, and hence its

- 21 -

ability to provide longer read lengths. The use of alkaline phosphatase in conjunction with apyrase and dATP $\alpha$ S (and optionally other dNTP $\alpha$ S's) has the unexpected additional benefit of abrogating these significant additional sources of inhibition (namely inhibition arising from the use of NTP $\alpha$ S together with apyrase, e.g. due to NDP $\alpha$ S and/or NMP $\alpha$ S produced by apyrase action).

Further benefits of using alkaline phosphatase may also be available. Thus, in certain conditions or circumstances, for example long-read sequencing (e.g. over 50 cycles), other, non-thio modified, NTPs added for incorporation, may also be a source of inhibition, albeit to a lesser extent than their  $\alpha$ -thio modified analogues. For example, preliminary results suggests that dATP may be a source of inhibition for the enzyme luciferase (see the Examples below). The action of apyrase on unincorporated NTPs results in the accumulation of degradation products, namely NDPs and NMPs (e.g. dADP, dAMP, dCTP, dCMP, dGDP, dGMP, dTDP, dTMP). In later cycles of sequencing, such products may inhibit enzymes used in the sequencing and/or PPi detection reactions e.g. apyrase and/or luciferase. Although any such inhibitory effects which may be obtained with other nucleotides will be lesser than that arising from use of NTP $\alpha$ S, they may still nonetheless contribute to a decreased efficiency of the system. Again, any potential problem caused by such inhibitory substances, may be removed or reduced by the action of alkaline phosphatase.

ATP is generated in the first step of the ELIDA reaction (catalysed by the enzyme ATP Sulphurylase) used as the preferred method for PPi detection. Such an ATP product (in particular, any such ATP not used in the subsequent luciferase reaction) may also act as a substrate for apyrase (or other nucleotide degrading enzyme) and hence may also lead to the generation of

- 22 -

inhibitory products (e.g. ADP and AMP) which may inhibit luciferase, and to a lesser extent, also apyrase. Again, any such inhibitory products may be reduced or removed by the action of alkaline phosphatase.

Finally, the use of alkaline phosphatase may further have a benefit in reducing any problems caused by kinase contamination. Kinases may be contaminants of enzyme preparations used in the sequencing and PPI detection reactions described herein. The action of kinases on apyrase degradation products (i.e. NDPs or NMPs), and ATP resulting from the ATP sulphurylase reaction in the ELIDA procedure may generate NTPs which may distort the primary sequencing reaction (i.e. may be incorporated by polymerase), and lead to non-synchronised extension. Alkaline phosphatase has a beneficial effect in degrading potential kinase substrates.

The sample nucleic acid (i.e. the target nucleic acid to be sequenced) may be any polynucleotide sequence it is desirable to obtain sequence information about. Thus, it may be any polynucleotide, or indeed oligonucleotide sequence. The nucleic acid may be DNA or RNA, and may be natural or synthetic. Thus, the target nucleic acid may be genomic DNA, or cDNA, or a PCR product or other amplicon etc. The target (sample) nucleic acid may be used in any convenient form, according to techniques known in the art e.g. isolated, cloned, amplified etc., and may be prepared for the sequencing reaction, as desired, according to techniques known in the art. The sample nucleic acid acts as a template for possible polymerase based extension of the primer and thus may conveniently be referred to as "template" or "nucleic acid template". The DNA may also be single or double-stranded - whilst a single-stranded DNA template has traditionally been used in sequencing reactions, or indeed in any primer-extension reaction, it is possible to use a double-stranded template; strand

- 23 -

displacement, or a localised opening-up of the two DNA strands may take place to allow primer hybridisation and polymerase action to occur.

In the polymerase reaction, any convenient polymerase enzyme may be used according to choice, as will be described in more detail below. In the case of a RNA template, such a polymerase enzyme may be a reverse transcriptase enzyme.

In order to repeat the method cyclically and thereby sequence the sample DNA and, also to aid separation of a single stranded sample DNA from its complementary strand, the sample DNA may optionally be immobilised or provided with means for attachment to a solid support.

Moreover, the amount of DNA present in a sample to be analysed may be small and it may therefore be desirable to amplify the DNA prior to sequencing. As mentioned above the sample DNA may thus be an amplicon.

Any desired method of *in vitro* or *in vivo* amplification may be used, e.g. PCR (or a variant or modification thereof) or Self Sustained Sequence Replication (3SR) or the ligase chain reaction (LCR) etc. Whichever method of amplification is used, it may be convenient to immobilise the amplified DNA, or provide it with means for attachment to a solid support. For example, a PCR primer may be immobilised or be provided with means for attachment to a solid support.

Immobilisation of the amplified DNA may take place as part of PCR amplification itself, as where one or more primers are attached to a support, or alternatively one or more of the PCR primers may carry a functional group permitting subsequent immobilisation, e.g. a biotin or thiol group. Immobilisation by the 5' end of a primer allows the strand of DNA emanating from that primer to be attached to a solid support and have its 3' end remote from the support and available for subsequent hybridisation with the extension primer and chain



- 24 -

extension by polymerase.

The solid support may conveniently take the form of microtitre wells. However, any solid support may conveniently be used, including any of the vast number described in the art, e.g. for separation/immobilisation reactions or solid phase assays. Thus, the support may also comprise particles (e.g. beads), fibres or capillaries made, for example, of agarose, cellulose, alginate, Teflon or polystyrene. Magnetic particles, e.g. the superparamagnetic beads produced by Dynal AS (Oslo, Norway) also may be used as a support.

The solid support may carry functional groups such as hydroxyl, carboxyl, aldehyde or amino groups, or other moieties such as avidin or streptavidin, for the attachment of nucleic acid molecules e.g. primers. These may in general be provided by treating the support to provide a surface coating of a polymer carrying one of such functional groups, e.g. polyurethane together with a polyglycol to provide hydroxyl groups, or a cellulose derivative to provide hydroxyl groups, a polymer or copolymer of acrylic acid or methacrylic acid to provide carboxyl groups or an aminoalkylated polymer to provide amino groups. US Patent No. 4654267 describes the introduction of many such surface coatings.

If desired, the sample may be washed after a certain number of reaction cycles e.g. 15-25, according to techniques well known in the art. Washing may be facilitated by immobilising the sample on a solid surface. Using a nucleotide-degrading enzyme, combined with Rp isomer elimination (particularly using alkaline phosphatase) however, means that washing is not absolutely necessary.

The assay technique is very simple and rapid, thus making it easy to automate by using a robot apparatus where a large number of samples may be rapidly analysed. Since the preferred detection and quantification is based on a luminometric reaction, this can be easily

- 25 -

followed spectrophotometrically. The use of luminometers is well known in the art and described in the literature.

The method of the present invention is particularly suited for use in an array format, wherein samples are distributed over a surface, for example a microfabricated chip, and thereby an ordered set of samples may be immobilized in a 2-dimensional format. Many samples can thereby be analysed in parallel. Using the method of the invention, many immobilized templates may be analysed in this way by allowing the solution containing the enzymes and one nucleotide to flow over the surface and then detecting the signal produced for each sample. This procedure can then be repeated. Alternatively, several different oligonucleotides complementary to the template may be distributed over the surface followed by hybridization of the template. Incorporation of nucleotides may be monitored for each oligonucleotide by the signal produced using the various oligonucleotides as primer. By combining the signals from different areas of the surface, sequence-based analyses may be performed by four cycles of polymerase reactions using the various nucleotides.

The length of the extension primer is not critical and can be according to choice. It will be clear to persons skilled in the art that the size of the extension primer and the stability of hybridisation will be dependent to some degree on the ratio of A-T to C-G base pairings, since more hydrogen bonding is available in a C-G pairing. Also, the skilled person will consider the degree of sequence identity between the extension primer to other parts of the template sequence and choose the degree of stringency accordingly. Guidance for such routine experimentation can be found in the literature, for example, Molecular Cloning: a laboratory manual by Sambrook, J., Fritsch E.F. and Maniatis, T. (1989). It may be advantageous to ensure

- 26 -

that the sequencing primer hybridises at least one base inside from the 3' end of the template to eliminate blunt-ended DNA polymerase activity. If separate aliquots are used (i.e. 4 aliquots, one for each base), the extension primer is preferably added before the sample is divided into four aliquots although it may be added separately to each aliquot.

Alternatively, a primer with a phosphorylated 5'-end, containing a loop and annealing back on itself and the 3'-end of the single stranded template can be used. If the 3'-end of the template has the sequence region denoted T (template), the primer has the following sequence starting from the 5'-end; P-L-P'-T', where P is primer specific (5 to 30 nucleotides), L is loop (preferably 4 to 10 nucleotides), P' is complementary to P (preferably 5 and 30 nucleotides) and T' is complementary to the template sequence in the 3'-end (T) (at least 4 nucleotides). This primer can then be ligated to the single stranded template using T4 DNA ligase or a similar enzyme. This provides a covalent link between the template and the primer, thus avoiding the possibility that the hybridised primer is washed away during the protocol.

In the polymerase reaction, any convenient polymerase enzyme may be used according to choice, e.g. T7 polymerase, Klenow or Sequenase Ver. 2.0 (USB U.S.A.). Any suitable polymerase may conveniently be used and many are known in the art and reported in the literature. However, it is known that many polymerases have a proof-reading or error checking ability and that 3' ends available for chain extension are sometimes digested by one or more nucleotides. If such digestion occurs in the method according to the invention the level of background noise increases. In order to avoid this problem, a nonproof-reading polymerase, e.g. exonuclease deficient (exo<sup>-</sup>) Klenow polymerase may be used and this is preferred according to the present

- 27 -

invention. Alternatively, substances which suppress 3' digestion by polymerase, such as fluoride ions or nucleotide monophosphates, may be used. The precise reaction conditions, concentrations of reactants etc. may readily be determined for each system according to choice. However, it may be advantageous to use an excess of polymerase over primer/template to ensure that all free 3' ends are extended.

In the method of the invention, it is preferred to use a DNA polymerase with high efficiency in each extension step due to the rapid increase of background signal which may take place if templates which are not fully extended accumulate. A high fidelity in each step is also desired, which can be achieved by using polymerases with exonuclease activity. However, this has the disadvantage mentioned above that degradation of the extending strand may occur. Although the exonuclease activity of the Klenow polymerase is low, the 3' end of the primer may be degraded with longer incubations in the absence of nucleotides. An induced-fit binding mechanism in the polymerisation step selects very efficiently for binding of the correct dNTP with a net contribution towards fidelity of  $10^5$ - $10^6$ . Exonuclease-deficient polymerases, such as (exo<sup>-</sup>) Klenow or Sequenase 2.0, catalysed incorporation of a nucleotide which was only observed when the complementary dNTP was present, confirming a high fidelity of these enzymes even in the absence of proof-reading exonuclease activity. The main advantage of using (exo<sup>-</sup>) Klenow DNA polymerase over Sequenase 2.0 is its lower  $K_m$  for nucleotides, allowing a high rate of nucleotide incorporation even at low nucleotide concentrations. As mentioned above, it is also possible to replace all dNTPs with nucleotide analogues or non-natural nucleotides such as dNTP $\alpha$ S, and such analogues may be preferable for use with a DNA polymerase having exonuclease activity.

- 28 -

In certain circumstances, e.g. with longer sample templates, it may be advantageous to use a polymerase which has a lower  $K_m$  for incorporation of the correct (matched) nucleotide, than for the incorrect (mismatched) nucleotide. This may improve the accuracy and efficiency of the method. Suitable such polymerase enzymes include the  $\alpha$ -polymerase of *Drosophila*.

In many diagnostic applications, for example genetic testing for carriers of inherited disease, the sample will contain heterozygous material, that is half the DNA will have one nucleotide at the target position and the other half will have another nucleotide. Thus if four aliquots (i.e. four parallel reactions of the same sample) are used in an embodiment according to the invention, two will show a negative signal and two will show half the positive signal. It will be seen therefore that it is desirable to quantitatively determine the amount of signal detected in each sample. Also, it will be appreciated that if two or more of the same base are adjacent the 3'-end of the primer a larger signal will be produced. In the case of a homozygous sample it will be clear that there will be three negative and one positive signal when the sample is divided into four parallel reactions.

In carrying out the method of the invention, any possible contamination of the reagents e.g. the NTP solutions, by PPi is undesirable and may readily be avoided by including a pyrophosphatase, preferably in low amounts, in the reagent solutions. Indeed, it is desirable to avoid contamination of any sort and the use of high purity or carefully purified reagents is preferred, e.g. to avoid contamination by kinases.

Reaction efficiency may be improved by including  $Mg^{2+}$  ions in the reagent (NTP and/or polymerase) solutions.

A potential problem which has previously been observed with PPi-based sequencing method arises when

- 29 -

the DNA to be sequenced has a number of identical adjacent bases, especially three or more of the same. Further, false signals may occur due to mispriming, i.e. hybridisation of the primer not to its targeted complement within the target DNA sequence but to another region, which will result in generation of "incorporation signals" which do not reflect the identity of the target sequence. As described in WO00/43540, this may be avoided by the use of single-stranded nucleic acid binding proteins in the reaction mixture after annealing of the primer to the template.

Thus, a further preferred feature of the invention is the use of a single-stranded nucleic acid binding protein, which is included during the polymerase reaction step after primer annealing.

As mentioned above, the benefits of the present invention arise from the elimination (e.g. removal and/or exclusion) of inhibiting substances from the reaction mixture.

Thus, viewed from a different aspect, the present invention provides a method of decreasing the inhibition of polymerase in a PPi-based sequencing procedure which uses at least one NTP $\alpha$ S, said method comprising eliminating the Rp isomer of NTP $\alpha$ S and/or the degradation products of said NTP $\alpha$ S from the sequencing reaction mixture (i.e. the template, primer, polymerase and/or nucleotide mix).

Further, since a similar inhibiting effect has been observed for the apyrase enzyme, which may be used to degrade un-incorporated nucleotides, the present invention also provides a method of decreasing the inhibition of apyrase when used in a PPi-based sequencing procedure which uses at least one NTP $\alpha$ S, said method comprising eliminating the Rp isomer of NTP $\alpha$ S and/or the degradation products of said NTP $\alpha$ S from the sequencing reaction.

As further mentioned above, the luciferase enzyme

- 30 -

which is preferably used in PPI detection may be inhibited by various inhibitory substances (e.g. Rp isomer and/or degradation products) and these may advantageously be removed by the action of alkaline phosphatase. Accordingly, in a further aspect the present invention provides a method of decreasing the inhibition of luciferase when used as a detection enzyme in a PPI-based sequencing procedure, said method comprising including alkaline phosphatase in the sequencing and/or PPI detection reaction mixture.

Typically, in certain embodiments of the invention NTP $\alpha$ S will be ATP $\alpha$ S (e.g. dATP $\alpha$ S or ddATP $\alpha$ S).

The invention also comprises kits for use in methods of the invention which will normally include at least the following components:

- (a) a polymerase;
- (b) means for detecting pyrophosphate release (preferably enzyme means, and most preferably luciferase and ATP sulphurylase, e.g. the reaction components of an ELIDA assay (see above);
- (c) optionally a nucleotide-degrading enzyme (preferably apyrase);
- (d) alkaline phosphatase
- (e) one or more nucleotides, preferably deoxynucleotides, including, in place of an adenine nucleotide (e.g. dATP), an  $\alpha$ -thiotriphosphate analogue of said nucleotide (e.g. dATP $\alpha$ S);
- (f) optionally, a test specific primer which hybridises to sample nucleic acid so that the target position is in close proximity to the 3' end of the primer;

- 31 -

A further embodiment of the kit of the invention will normally include at least the following components:

- (a) a polymerase;
- (b) means for detecting pyrophosphate release (preferably enzyme means, and most preferably luciferase and ATP sulphurylase, e.g. the reaction components of an ELIDA assay (see above);
- (c) optionally a nucleotide-degrading enzyme (preferably apyrase);
- (d) the Sp isomer of an  $\alpha$ -thiotriphosphate analogue of an adenine nucleotide (preferably dATP $\alpha$ S) and optionally one or more Sp isomers of thymine, cytosine or guanine nucleotides (preferably dGTP $\alpha$ S, dCTP $\alpha$ S or dGTP $\alpha$ ); and
- (f) optionally, a test specific primer which hybridises to sample nucleic acid DNA so that the target position is in close proximity to the 3' end of the primer;
- (g) optionally, one or more unmodified thymine, cytosine or guanine nucleotides (preferably dTTP, dCTP, dGTP);
- (h) optionally, alkaline phosphatase.

The benefits of the present invention in reducing the effects of inhibitory substances in PPi-based sequencing, improve the efficiency and reliability of the method and extend the applications in which the method can be used. Thus, techniques requiring longer reads such as genome re-sequencing, comparative EST sequencing, microbial typing and confirmatory sequencing



- 32 -

may all be conducted by PPi-based sequencing according to the present invention.

The primer, if present in the kits, is designed so that it binds to the template nucleic acid with its 3' end in close proximity to the target nucleotide. By close proximity it is meant 1 to 30 nucleotides, preferably 1 to 20 nucleotides, more preferably 1 to 10, most preferably 1 to 5 nucleotides from the target nucleotide.

The invention will now be described by way of non-limiting examples with reference to the drawings in which:

Figure 1 shows a trace (light intensity v nucleotide addition) obtained from simulation of the effect of dATP on luciferase and apyrase, using 2  $\mu$ M ATP in the nucleotide solution. In this experiment, 100 ng luciferase, 50 mU apyrase, 0.1 M Tris-acetate (pH 7.75), 0.5 mM EDTA, 5 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 0.4 mg/ml polyvinylpyrrolidone (360 000), and 100  $\mu$ g/ml D-luciferin (BioThema) were used. 200  $\mu$ l of 0.7 mM nucleotide (dATP) containing 2  $\mu$ M ATP is dispensed into the reaction mixture as described in Example 1. The height of the ascending curve demonstrates luciferase activity and the slope of the descending curve demonstrates the apyrase activity. The PPi was detected in real time.

Figure 2 shows a trace (light intensity v nucleotide addition) obtained from simulation of the effect of dATP $\alpha$ S on luciferase and apyrase, using 2  $\mu$ M ATP in the nucleotide solution. In this experiment, 100 ng luciferase, 50 mU apyrase, 0.1 M Tris-acetate (pH 7.75), 0.5 mM EDTA, 5 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 0.4 mg/ml polyvinylpyrrolidone (360 000), and 100  $\mu$ g/ml D-luciferin (BioThema) were used. 200  $\mu$ l of 0.7 mM nucleotide (dATP $\alpha$ S) containing 2  $\mu$ M ATP is dispensed to

- 33 -

the reaction mixture as described in Example 1. The height of ascending curve demonstrates luciferase activity and the slope of descending curve demonstrates the apyrase activity.

Figure 3 shows a trace (light intensity v nucleotide addition) obtained from simulation of the effect of dCTP on luciferase and apyrase, using 2  $\mu$ M ATP in the nucleotide solution. In this experiment, 100 ng luciferase, 50 mU apyrase, 0.1 M Tris-acetate (pH 7.75), 0.5 mM EDTA, 5 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 0.4 mg/ml polyvinylpyrrolidone (360 000), and 100  $\mu$ g/ml D-luciferin (BioThema) were used. 200  $\mu$ l of 0.2 mM nucleotide (dCTP) containing 2  $\mu$ M ATP is dispensed to the reaction mixture as described in Example 1. The height of ascending curve demonstrates luciferase activity and the slope of descending curve demonstrates the apyrase activity.

Figure 4 shows a trace (light intensity v nucleotide addition) obtained from simulation of the effect of dGTP on luciferase and apyrase, using 2  $\mu$ M ATP in the nucleotide solution. In this experiment, 100 ng luciferase, 50 mU apyrase, 0.1 M Tris-acetate (pH 7.75), 0.5 mM EDTA, 5 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 0.4 mg/ml polyvinylpyrrolidone (360 000), and 100  $\mu$ g/ml D-luciferin (BioThema) were used. 200  $\mu$ l of 0.16 mM nucleotide (dGTP) containing 2  $\mu$ M ATP is dispensed to the reaction mixture as described in Example 1. The height of ascending curve demonstrates luciferase activity and the slope of descending curve demonstrates the apyrase activity.

Figure 5 shows a trace (light intensity v nucleotide addition) obtained from simulation of the effect of dTTP on luciferase and apyrase, using 2  $\mu$ M ATP in the nucleotide solution. In this experiment, 100 ng luciferase, 50 mU apyrase, 0.1 M Tris-acetate (pH 7.75),

- 34 -

0.5 mM EDTA, 5 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 0.4 mg/ml polyvinylpyrrolidone (360 000), and 100 µg/ml D-luciferin (BioThema) were used. 200 µl of 0.8 mM nucleotide (dTTP) containing 2 µM ATP is dispensed to the reaction mixture as described in Example 1. The height of ascending curve demonstrates luciferase activity and the slope of descending curve demonstrates the apyrase activity.

Figure 6 shows a trace (light intensity v nucleotide addition) obtained from simulation of the effect of dATPαS on luciferase and apyrase, using 2 µM ATP in the nucleotide solution, in the presence of alkaline phosphatase. In this experiment, 100 ng luciferase, 50 mU apyrase, 0.1 M Tris-acetate (pH 7.75), 0.5 mM EDTA, 5 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 0.4 mg/ml polyvinylpyrrolidone (360 000), 100 µg/ml D-luciferin (BioThema), and 1U alkaline phosphatase were used. 200 µl of 0.7 mM nucleotide (dATPαS) containing 2 µM ATP is dispensed to the reaction mixture as described in Example 1. The height of ascending curve demonstrates luciferase activity and the slope of descending curve demonstrates the apyrase activity.

Figure 7 shows three traces (light intensity v nucleotide addition) obtained in DNA sequencing reactions using 1 pmol of primed oligonucleotide template; the reaction system contained 10 U Klenow DNA polymerase, 25 mU ATP sulphurylase and 0.1 M Tris-acetate (pH 7.75), 0.5 mM EDTA, 5 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 0.4 mg/ml polyvinylpyrrolidone (360 000), and 100 µg/ml D-luciferin (BioThema); and A) 100 ng luciferase and 50 mU apyrase; B) 100 ng luciferase, 50 mU apyrase and 2 mU nucleoside diphosphate kinase; C) 100 ng luciferase, 50 mU apyrase, 2 mU nucleoside diphosphate kinase and 2 U alkaline phosphatase. The arrows on B indicates the

- 35 -

false signals appearing as a result of kinase activity in the system, which are removed by addition of alkaline phosphatase in the experiment shown in C.

Figure 8 presents traces (light intensity v nucleotide addition) showing the effects of pre-incubating a PPI-sequencing reaction mixture (in the absence of template) with varying amounts of "normal" (i.e. racemic) dATP $\alpha$ S (containing both Rp and Sp isomers) ((A) 0  $\mu$ l; (B) 5  $\mu$ l; (C) 10  $\mu$ l; (D) 20  $\mu$ l) or the pure Sp isomer of dATP $\alpha$ S ((E) 0  $\mu$ l; (F) 5  $\mu$ l; (G) 10  $\mu$ l; (H) 20  $\mu$ l) or the pure Rp isomer of dATP $\alpha$ S ((I) 0  $\mu$ l; (J) 5  $\mu$ l; (K) 10  $\mu$ l; (L) 20  $\mu$ l). After pre-incubation, template was added and the PPI-based sequencing reaction was performed as described in Example 2.

Figure 9 is a trace (light intensity v nucleotide addition) showing the results of a DNA sequencing reaction on a pUC19-derived template wherein the four different nucleotides are added stepwise to the template hybridized to a primer. A mixture of the Rp and Sp isomers of dATP $\alpha$ S is used. The template/primer was incubated with 10U (exo<sup>-</sup>) Klenow and 40 mU apyrase, and other components of the sequencing/PPI detection reaction as described in Example 3. The reaction was started by the addition of the first nucleotide, and the nucleotides were added in a stepwise fashion. The PPI released was detected in real time by luciferase.

Figure 10 is a trace (light intensity v nucleotide addition) showing the results of a DNA sequencing reaction on a pUC19-derived template wherein the four nucleotides are added stepwise to the template hybridized to the primer. Only the pure Sp isomer of d-ATP $\alpha$ S was used. The template/primer was incubated with 10U (exo<sup>-</sup>) Klenow and 40 MU apyrase and other components of the sequencing/PPI detection reaction as described in Example 3. The reaction was started by the addition of the first nucleotide, and nucleotides were then added in a stepwise fashion. The PPI released was detected in

- 36 -

real time by luciferase.

- 37 -

Example 1Investigation of the inhibitory effects of different nucleotides in PPI-based sequencing using an ELIDA detection reaction - the effects of alkaline phosphatase

In this Example a series of 5 experiments were carried out to examine the inhibitory effect of each of the 5 nucleotides (dATP, dATP $\alpha$ S, dCTP, dGTP and dTTP) on the enzymes apyrase and luciferase. A model system was used which simulated the effects of these enzymes in a PPI-based sequencing reaction. Each nucleotide was added to the luciferase/apyrase/substrate mixture, along with ATP (to provide a substrate for luciferase), and the effect of each nucleotide addition on the enzymes was monitored by observing the light signals generated from the luciferase reaction. A further two experiments demonstrate, firstly, that the inhibitory effects resulting from use of dATP $\alpha$ S may be abrogated in the model system by the use of alkaline phosphatase, and, secondly that alkaline phosphatase further has an additional benefit in counteracting the interfering effects of kinase enzymes on the reaction system (as demonstrated in a sequencing reaction).

The reactions were performed at room temperature in a volume of 50  $\mu$ l on an automated prototype Pyrosequencer instrument (kindly supplied by Pyrosequencing AB, Uppsala, Sweden). The reaction mixture contained: 50 mU apyrase (Sigma Chemical Co., USA), 100 ng purified luciferase (BioThema, Dalarö, Sweden), 0.1 M Tris-acetate (pH 7.75), 0.5 mM EDTA, 5 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 0.4 mg/ml polyvinylpyrrolidone (360 000), and 100  $\mu$ g/ml D-luciferin (BioThema). An ATP concentration of 2  $\mu$ M was added to different nucleotides (dATP, dATP $\alpha$ S, dCTP, dGTP and dTTP), as shown in Figures 1 to 5 to follow the

- 38 -

activity of luciferase and apyrase. For investigation of the effect of Shrimp Alkaline Phosphatase (Amersham Pharmacia Biotech, Uppsala, Sweden), two units of this enzyme were added to the reaction mixture - see further below. In total, 50 additions of nucleotide and ATP mixture were performed in 50 minutes. The output of light resulting from nucleotide incorporation was detected by a CCD-camera. The data was obtained in Excel format.

Seven experiments were performed. Experiments 1 to 5 investigated the effects of each of dATP, dATP $\alpha$ S, dCTP, dGTP and dTTP. Experiment 1 was performed using an admixture of 0.7 mM nucleotide dATP and 2  $\mu$ M ATP dispensed to the reaction mixture. The trace obtained from experiment 1 is presented as Figure 1. The height of the ascending curve demonstrates luciferase activity and the slope of the descending curve is indicative of or demonstrates apyrase activity. Note the relatively linear decrease in signal intensity, which is due to inhibition of luciferase, believed to be due to accumulation of degradation products of dATP (e.g. dAMP). It has earlier been reported that dAMP inhibits luciferase (Ford et al., 1998, Methods in Mol. Biol., 102: 3-20) and accordingly it is inferred from this that the inhibitory substance may be dAMP. Thus, luciferase inhibition is shown, when dATP is used, and this is believed to be due to the degradation products of dAMP (dATP and/or dADP).

Experiment 2 was performed using an admixture of 0.7 mM nucleotide dATP $\alpha$ S and 2  $\mu$ M ATP dispensed to the reaction mixture. The results are shown in Figure 2. The height of the ascending curve demonstrates luciferase activity and the slope of the descending curve demonstrates apyrase activity. Note the relatively linear decrease in signal intensity, which is due to inhibition of

- 39 -

luciferase, believed to be caused by accumulation of inhibitory products (e.g. dAMP $\alpha$ S). The decrease in signal intensity after 50 cycles is 50% less than the data obtained from degradation of dATP in Figure 1. Accordingly, it is inferred from this that just one isomer of dATP $\alpha$ S is inhibitory, and this is believed to be the Sp-isomer (or rather that the products of Sp isomer are inhibitory to luciferase). Apyrase is drastically inhibited in later cycles as seen by the wider peaks which are obtained. The inhibitory effect after each addition is believed to be due to accumulation of the Rp isomer of dATP $\alpha$ S, and also the product(s) of the degradation of the Sp isomer. It is believed that apyrase degrades only the Sp isomer and not the Rp isomer. It is speculated that the inactive isomer of dATP $\alpha$ S (Rp) is not recognised by luciferase since it is not degraded to form the inhibitory product.

Experiment 3 was performed using an admixture of 0.2 mM nucleotide dCTP and 2  $\mu$ M ATP dispensed to the reaction mixture. The results are shown in Figure 3. The height of the ascending curve demonstrates luciferase activity and the slope of the descending curve demonstrates apyrase activity. Note the relatively constant signal intensity even after 50 cycles which indicates that the products of this nucleotide do not inhibit luciferase or apyrase under the conditions used, and within the number of cycles shown.

Experiment 4 was performed using an admixture of 0.16 mM nucleotide dGTP and 2  $\mu$ M ATP dispensed to the reaction. The results obtained from experiment 4 are presented in Figure 4. The height of the ascending curve demonstrates luciferase activity and the slope of the descending curve demonstrates the apyrase activity. Note the relatively constant signal intensity even after 50 cycles which indicate that the products of this



- 40 -

nucleotide do not inhibit luciferase or apyrase under the conditions used, and within the number of cycles shown.

Experiment 5 was performed using an admixture of 0.8 mM nucleotide dTTP and 2  $\mu$ M ATP dispensed to the reaction. The results are shown in Figure 5. The height of ascending curve demonstrates luciferase activity and the slope of descending curve demonstrates the apyrase activity. Note the relatively constant signal intensity even after 50 cycles which indicate that the products of this nucleotide do not inhibit luciferase or apyrase, under the conditions used, and within the number of cycles shown.

Experiment 6 investigated the reduction in inhibition effected by the addition of alkaline phosphatase to the reaction mixture. This experiment was performed similarly as for Experiment 2, using an admixture of 0.7 mM nucleotide dATP $\alpha$ S and 2  $\mu$ M ATP dispensed to the reaction mixture. The results are shown in Figure 6. The height of the generated peak demonstrates luciferase activity and the slope of the descending curve demonstrates the apyrase activity. Note the efficiency of nucleotide degradation as compared with Figure 2. It can be seen that the signal width remains constant even after 50 cycles indicating lack of apyrase inhibition, and hence the high efficiency of alkaline phosphatase in preventing this inhibition (believed to be due to its action in degrading both isomers (i.e. the Rp and Sp isomers) of dATP $\alpha$ S). No decrease in signal intensity is observed, indicating no inhibition of luciferase. This is believed to be due to the effect of alkaline phosphatase in degrading the products of dATP $\alpha$ S (thus removing the inhibitory elements). The experimental conditions for Figures 2 and 6 are the same except for addition of 2 U AP in the reaction mixture in Experiment

- 41 -

6. Thus, the results of Experiment 6 show that alkaline phosphatase can remove inhibitory substances from the reaction system, thus improving the performance of the luciferase and apyrase enzymes.

Experiment 7 investigated the inhibitory effects of nucleotides and kinases upon a sequencing reaction, and the reduction of these effects using alkaline phosphatase. For investigation of the effect of kinase, 2 mU of kinase (Sigma Chemicals) was added to the sequencing reaction mixture containing: 0.5 pmol primed synthetic template E3PN/NUSPT (Ronaghi et al. Science, 1998), 10 U exonuclease-deficient Klenow DNA polymerase (Amersham Pharmacia Biotech, Uppsala, Sweden), 40 mU apyrase (Sigma Chemical Co., USA), 100 ng purified luciferase (BioThema, Dalaro, Sweden), 15 mU of recombinant produced ATP sulfurylase, 0.1 M Tris-acetate (pH 7.75), 0.5 mM EDTA, 5 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 5  $\mu$ M adenosine 5-phosphosulfate (APS), 0.4 mg/ml polyvinylpyrrolidone (360 000), and 100  $\mu$ g/ml D-luciferin (BioThema). The sequencing procedure was carried out by stepwise elongation of the primer-strand upon sequential addition of Sp-dATP $\alpha$ S (Biolog Life Science, Bremen, DE), dCTP, dGTP, and dTTP (Amersham Pharmacia Biotech) and simultaneous degradation of nucleotides by apyrase. To study the effect of alkaline phosphatase on kinase activity, 2 mU of kinase enzyme and 2 U of alkaline phosphatase were added to the above admixture. The output of light resulting from nucleotide incorporation was detected by a CCD-camera. The data was obtained in Excel format. Figure 7a shows the results of the sequencing reaction in the absence of added kinase or alkaline phosphatase. Figure 7b shows a trace of the results of the sequencing reaction in the presence of 2 mU of nucleoside diphosphate kinase. The arrow indicates the false signals that are generated.

- 42 -

Thus, it can be seen that the addition of alkaline phosphatase to the polymerisation reaction mixture removes the false signals generated by kinase contamination, by removing the substrates for the kinase from solution.

### Example 2

#### Inhibitory effects of dATP $\alpha$ S on PPi-based sequencing

In this example, the inhibitory effect of dATP $\alpha$ S on a PPi-based sequencing reaction with ELIDA detection (known as Pyrosequencing<sup>TM</sup>) was investigated by preincubating the reaction mixture in the absence of template with varying amounts of the normal (i.e. racemic) dATP $\alpha$ S (containing both Rp and Sp isomers), or the pure Sp isomer of dATP $\alpha$ S. After preincubation, template was added and a normal "Pyrosequencing<sup>TM</sup>" reaction was carried out.

#### Materials and Methods

The "normal" dATP $\alpha$ S was taken from the PSQ 96 SNP Reagent kit supplied by Pyrosequencing AB (Uppsala, SE). The pure Rp and Sp isomers were purchased from Biolog Life Science (Bremen, Germany). The reaction mixture also included Enzyme mix (DNA polymerase, ATP-sulphurylase, apyrase and luciferase) and Substrate mix (luciferin and APS), from the PSQ 96 SNP Reagent kit supplied by Pyrosequencing AB. The template was an oligonucleotide (interactiva, Ulm, Germany) from which the following sequence could be read after annealing of a sequencing primer: CTAAAGGTGCACCATGACTGGGGTTACAGTCATC.

The pure isomer samples were diluted to the same concentration as the dATP $\alpha$ S in the PSQ 96 SNP Reagent kit. For preincubation, 0, 5, 10 or 20  $\mu$ l of each A-

- 43 -

sample (normal dATP $\alpha$ S or the Rp or Sp isomer) was added to a mixture containing 5  $\mu$ l Enzyme mix, 5  $\mu$ l Substrate mix in a total volume of 45  $\mu$ l. The reaction was carried out in the 96 well plate delivered with the PSQ 96 SNP Reagent kit. The reaction was incubated at room temperature, in the dark for 10 minutes. After this, 1.5 pmol in 5  $\mu$ l annealed oligonucleotide template was added to each well and the plate was transferred from the PSQ 96 instrument, where a normal "Pyrosequencing<sup>TM</sup>" reaction was carried out (omitting addition of Enzyme and Substrate mixes).

The results are shown in Figure 8. Using normal (i.e. racemic) dATP $\alpha$ S, it can be seen that pre-incubation with 5  $\mu$ l (Fig. 8B) 10  $\mu$ l (Fig. 8C) or 20  $\mu$ l (Fig. 8D) results in progressively severe inhibition of the sequencing reaction (as compared with 0  $\mu$ l of dATP $\alpha$ S added (Fig. 8A)). Some loss of signal intensity (peak height) and more significantly, loss of signal definition (e.g. broader, less defined and less clean signals) can be seen. Thus, it can be seen that apyrase inhibition is particularly occurring. Even more pronounced inhibition is seen when the Rp isomer is used (see Figure 8J). This effect was significantly reduced when using the pure Sp isomer in place of racemic dATP $\alpha$ S or Rp isomer of dATP $\alpha$ S, as seen from Fig. 8F (5  $\mu$ l), Fig. 8G (10  $\mu$ l) and Fig. 8H (20  $\mu$ l) (compared to 0  $\mu$ l (Fig. 8E)). The pure Rp isomer has a severe effect on all enzymes involved in the ELIDA detection (Figures 8I and 8J). Apyrase inhibition is indicated by increased signal width. This can clearly be seen in figures 8J to 8L, most clearly it is demonstrated on figure 8K. Luciferase inhibition is indicated by decreased signal intensity. As can be seen on figure 8K, the signal intensity is decreased in comparison with figure 8I, indicating that luciferase is being inhibited.

- 44 -

A similar experiment was also performed where the preincubation was carried out comparing "normal" dATP $\alpha$ S with dCTP, dGTP or dTTP, all from a prototype to the PSQ 96 SNP Reagent kit. As control, only 1xTE (dilution buffer for the dNTPs) was added. The template was the same as above, but 2 pmol were added to each well.

The results obtained (not shown) clearly showed no negative effect from preincubation with dCTP, dGTP or dTTP, but confirmed a severe negative effect from dATP $\alpha$ S, mainly on apyrase.

### Example 3

#### PPi-based DNA sequencing with and without Rp-dATP $\alpha$ S

In this Example, two sequencing experiments were performed using the same template (a PCR product of the standard cloning plasmid pUC19), with and without the Rp isomer of dATP $\alpha$ S (i.e. using firstly a racemic mixture of dATP $\alpha$ S, and secondly pure Sp-dATP $\alpha$ S).

#### Single stranded template preparation

The standard plasmid pUC19 was used to generate the template by PCR. In brief, PCR primers GGGATCATGTAACTCGCCTTGA (Upper primer, biotinylated position 1345) and CGGGAGGGCTTACCATCTGG (lower primer, position 1648), (where positions 1648-1345 = 303 bp) were used in a PCR reaction on pUC19 to generate a fragment of 303 bp in length. Fifty microliters of biotinylated PCR product was immobilized onto 20  $\mu$ l streptavidin-coated super paramagnetic beads (Dynabeads M-280-streptavidin, Dynal AS, Oslo, Norway) by incubation at 43°C for 30 minutes. Single-stranded DNA was obtained by incubating the immobilized PCR product in 5  $\mu$ l 0.1 M NaOH for 4 minutes. The immobilized

- 45 -

strand was resolved in 8  $\mu$ l H<sub>2</sub>O plus 1  $\mu$ l annealing buffer (100 mM Tris-Ac<sub>2</sub> (pH 7.75), 20 mM MgAc<sub>2</sub>). Single-stranded DNA corresponding to 50  $\mu$ l PCR product was hybridized to 10 pmol sequencing primer (TCAGCAATAAACCAGCCAGCC) at 70°C for 3 minutes followed by incubation at room temperature for 5 minutes. (After annealing of the sequencing primer, the length of single-stranded template remaining is 211 bp.) The primed PCR product was added to the Pyrosequencing™ reaction mixture containing: 0.1 M Tris-Ac<sub>2</sub> (pH 7.75), 0.05% Tween 20, 10 U exonuclease deficient (exo<sup>-</sup>) Klenow DNA polymerase, 40 mU apyrase (Sigma Chemical Co. St. Louis, Mo, USA), 0.8  $\mu$ g purified luciferase (BioThema, Dalarö, Sweden), 15 mU recombinant ATP sulfurylase (Karamohamed et al., 1999), 0.5  $\mu$ g single-stranded DNA binding protein (Amersham Pharmacia Biotech., Uppsala, Sweden), 0.5 mM EDTA, 5 mM MgAc<sub>2</sub>, 0.1% bovine serum albumin (BioThema), 1 mM dithiothreitol, 5  $\mu$ M adenosine 5'-phosphosulfate (Sigma Chem. Co.), 0.4 mg/ml polyvinylpyrrolidone (360 000), and 100  $\mu$ g/ml D-luciferin (BioThema) in a total volume of 50  $\mu$ l.

#### Pyrosequencing™

Pyrosequencing™ was performed at room temperature on an automated Pyrosequencer prototype model (Pyrosequencing AB, Uppsala, Sweden; [www.pyrosequencing.com](http://www.pyrosequencing.com)) at a dispensing pressure of 600 mbar with 8 msec open time and 60 sec cycle time. The sequencing procedure was carried out by stepwise elongation of the primer-strand upon cyclic dispensation of the different deoxynucleoside triphosphates (Amersham Pharmacia Biotech). In one experiment (shown in Figure 9), a racemic mixture of dATP $\alpha$ S ("normal" dATP $\alpha$ S) is used, (along with dCTP, dGTP and dTTP) and in the second experiment pure Sp isomer of dATP $\alpha$ S is used. The output of light resulting from nucleotide incorporation was

- 46 -

detected by a photomultiplier. The data was obtained in Microsoft Excel and is shown in Figures 9 and 10.

Looking at Figure 9 it can be seen that signal quality gradually deteriorates with repeated nucleotide addition, leading eventually to loss of readable signal (parts 4-6 of Fig. 9). In Figure 10, it will be seen that with pure Sp isomer, signal quality is maintained for longer, and a longer read-length is obtained before signal quality deteriorates below readable.

Claims

1. A method of identifying a base at a target position in a sample nucleic acid sequence, said method comprising:

subjecting a primer hybridised to said sample nucleic acid immediately adjacent to the target position, to a polymerase primer extension reaction in the presence of a nucleotide, whereby the nucleotide will only become incorporated if it is complementary to the base in the target position, and determining whether or not said nucleotide is incorporated by detecting whether PPi is released, the identity of the target base being determined from the identity of any nucleotide incorporated,

wherein, where said nucleotide comprises an adenine base, an  $\alpha$ -thio triphosphate analogue of said nucleotide is used, and the Rp isomer of said analogue and/or the degradation products of said analogue are eliminated from the polymerase reaction step.

2. A method as claimed in claim 1 wherein the Rp isomer is eliminated by using a preparation of an  $\alpha$ -thio triphosphate analogue of an adenine nucleotide which contains only the Sp isomer thereof.

3. A method as claimed in claim 2 wherein the Rp isomer is eliminated by using a preparation of dATP $\alpha$ S or dd ATP $\alpha$ S which contains only the Sp isomer thereof.

4. A method as claimed in claim 1 wherein the Rp isomer and/or the degradation products of said analogue are eliminated by enzymic degradation.

5. A method as claimed in claim 4 wherein alkaline phosphatase is included in or added to the polymerase reaction mixture.



- 48 -

6. A method as claimed in any preceding claim wherein a nucleotide degrading enzyme is present during or after the polymerase reaction step.

7. A method as claimed in claim 6 wherein the nucleotide degrading enzyme is apyrase..

8. A method as claimed in claim 2 wherein alkaline phosphatase is included in or added to the polymerase reaction mixture.

9. A method as claimed in any one of claims 1 to 8 wherein the primer extension reaction is repeated in the presence of further nucleotides.

10. A method of decreasing the inhibition of apyrase when used in a PPi-based sequencing procedure which uses at least one NTP $\alpha$ S, said method comprising eliminating the Rp isomer of NTP $\alpha$ S and/or the degradation products of said analogue from the sequencing reaction.

11. A method of decreasing the inhibition of luciferase when used in a PPi-based sequencing procedure which uses at least one NTP $\alpha$ S, said method comprising eliminating the Rp isomer of NTP $\alpha$ S and/or the degradation products of said analogue from the sequencing reaction.

12. A kit for use in a method of identifying a base at a target position in a nucleic acid which comprises:

- (a) a polymerase;
- (b) means for detecting pyrophosphate release;
- (c) optionally a nucleotide-degrading enzyme;
- (d) alkaline phosphatase

- 49 -

- (e) one or more nucleotides, preferably deoxynucleotides, including, in place of an adenine nucleotide, an  $\alpha$ -thiotriphosphate analogue of said nucleotide;
- (f) optionally, a test specific primer which hybridises to sample nucleic acid so that the target position is in close proximity to the 3' end of the primer;

13. A kit for use in a method of identifying a base at a target position in a nucleic acid which comprises:

- (a) a polymerase;
- (b) means for detecting pyrophosphate release;
- (c) optionally a nucleotide-degrading enzyme;
- (d) the Sp isomer of an  $\alpha$ -thiotriphosphate analogue of an adenine nucleotide and optionally one or more Sp isomers of thymine, cytosine or guanine nucleotides; and
- (f) optionally, a test specific primer which hybridises to sample nucleic acid DNA so that the target position is in close proximity to the 3' end of the primer;
- (g) optionally, one or more unmodified thymine, cytosine or guanine nucleotides;
- (h) optionally, alkaline phosphatase.

1 / 22

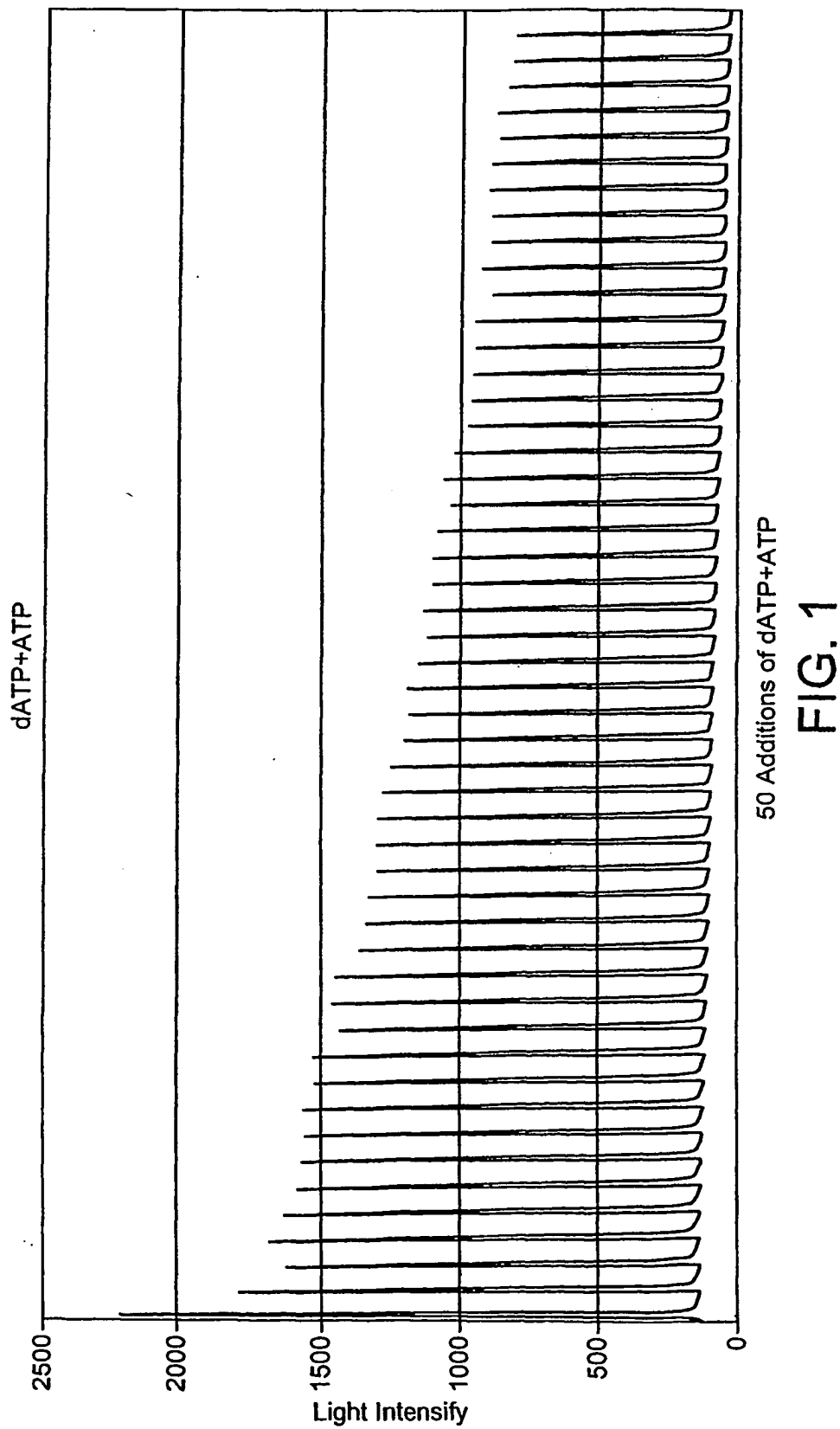
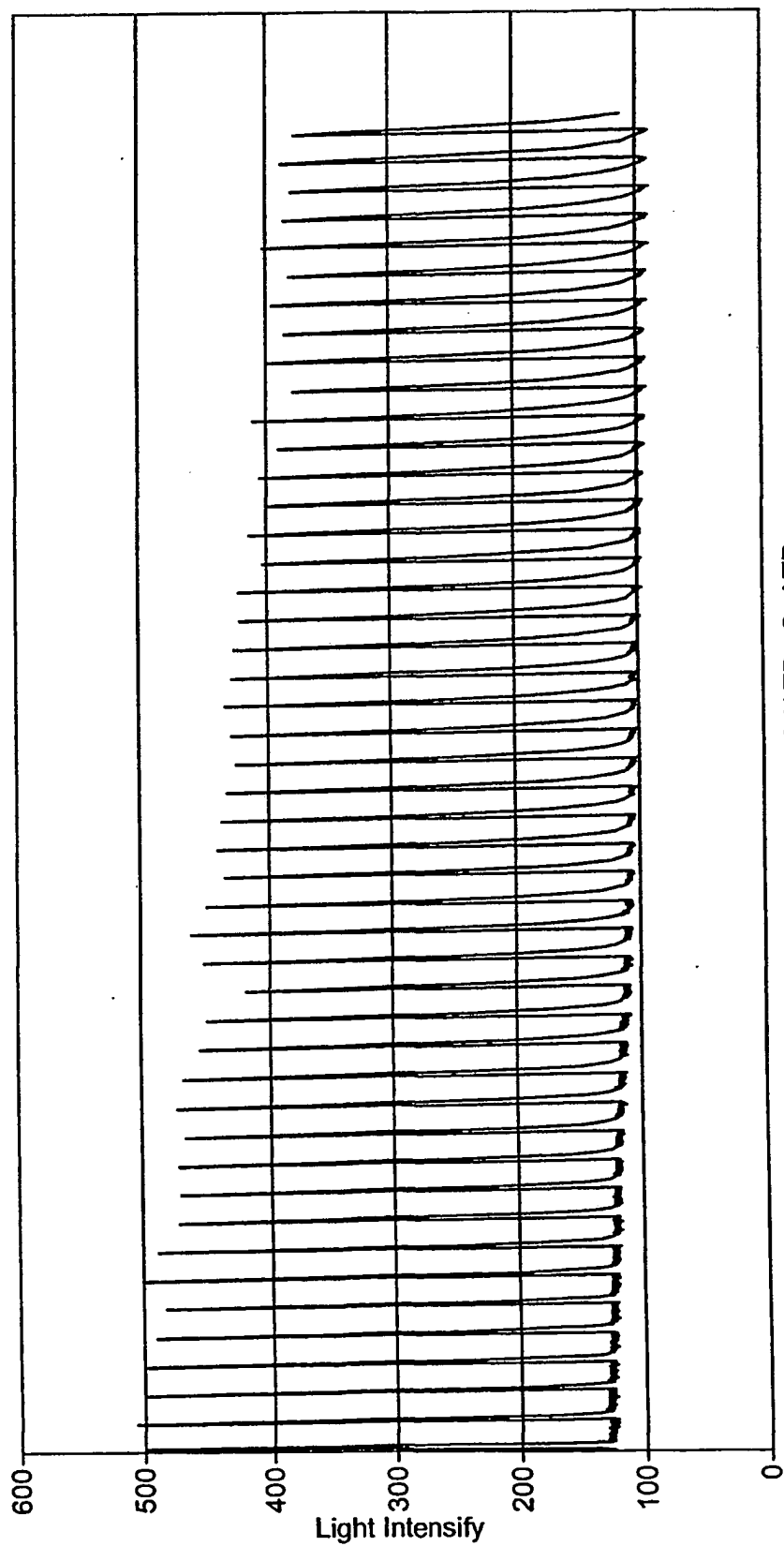


FIG. 1

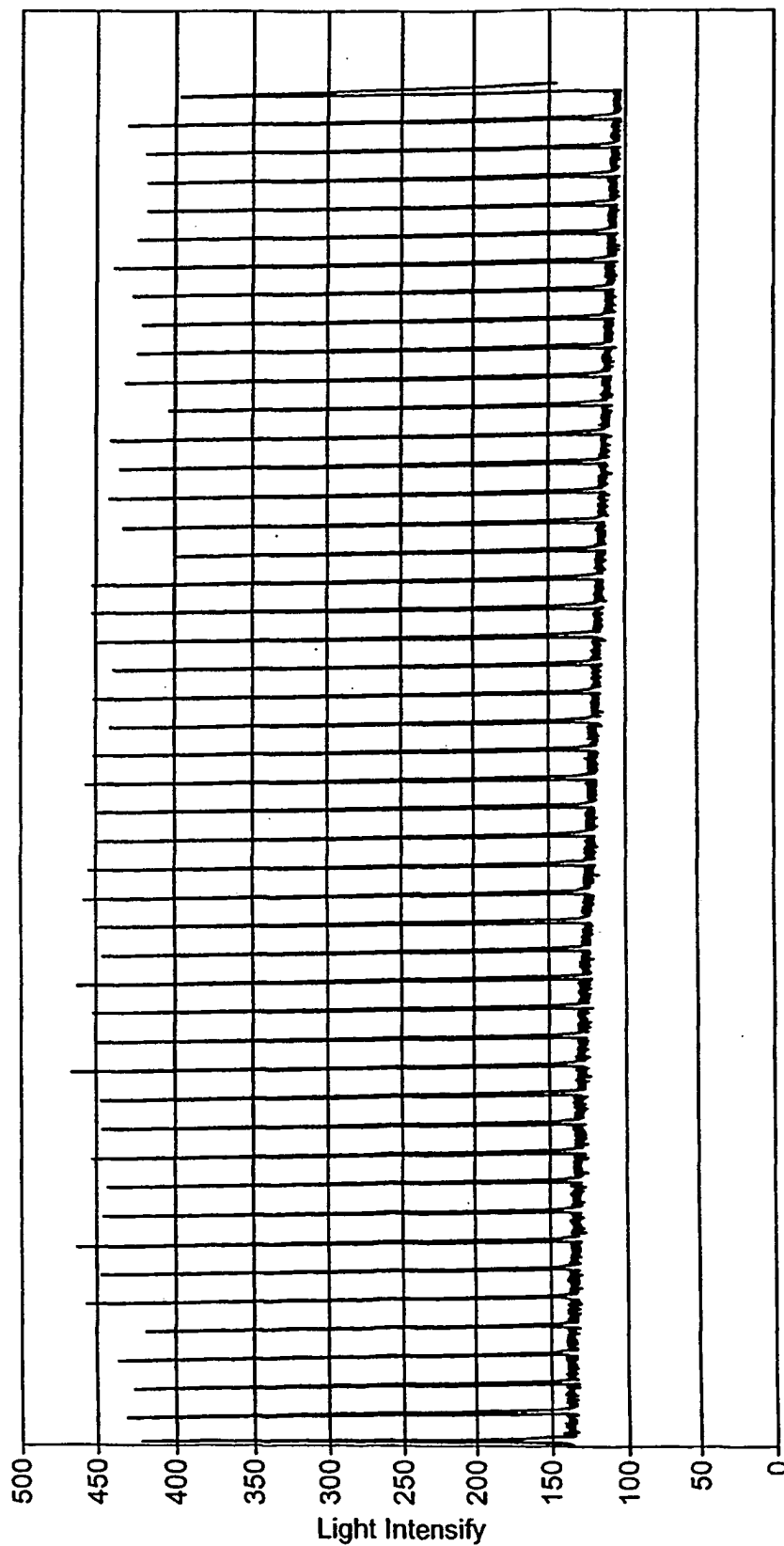
2 / 22



50 Additions of dATPα S+ATP

FIG. 2

3 / 22



50 Additions of dCTP+ATP

FIG. 3

4 / 22

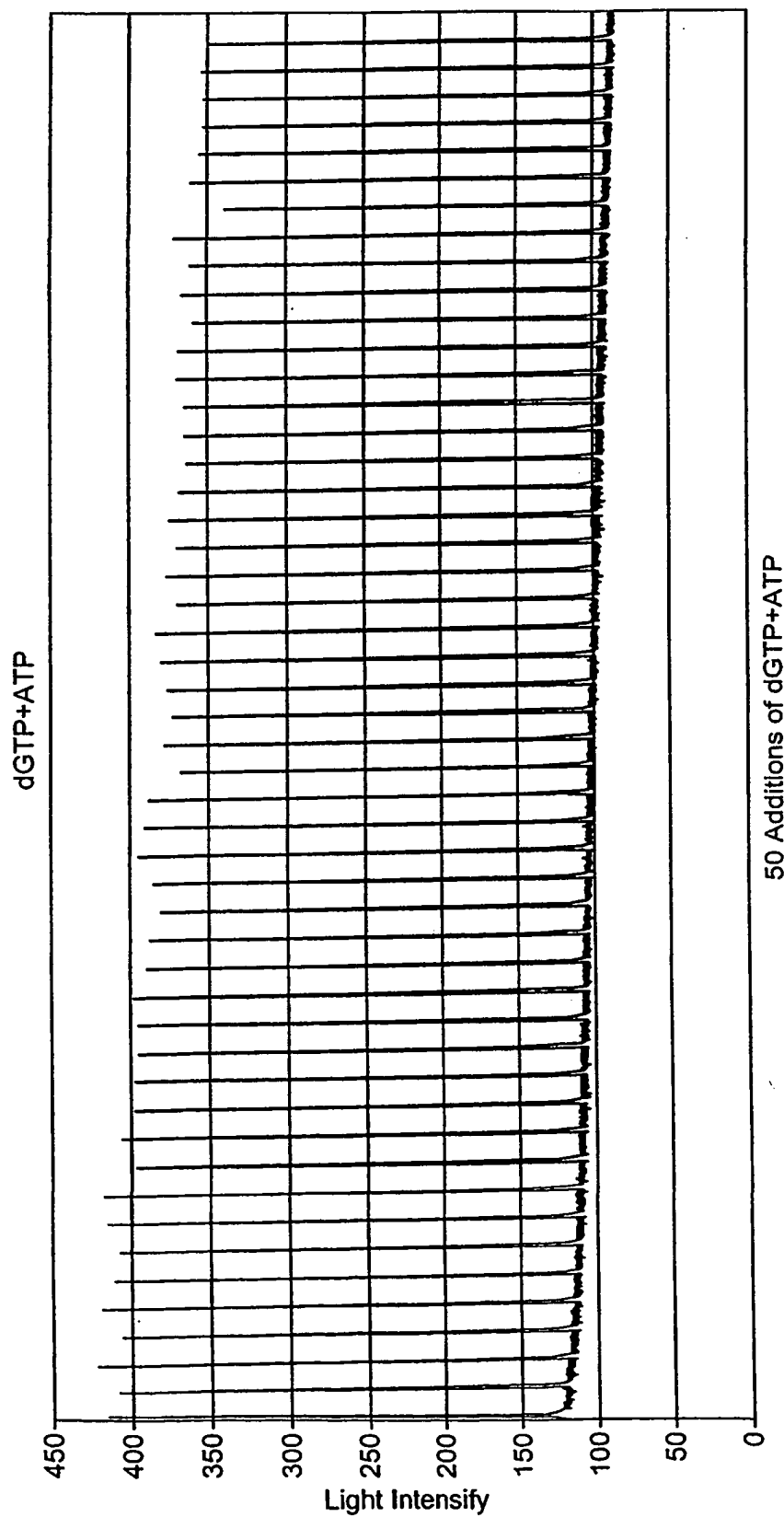
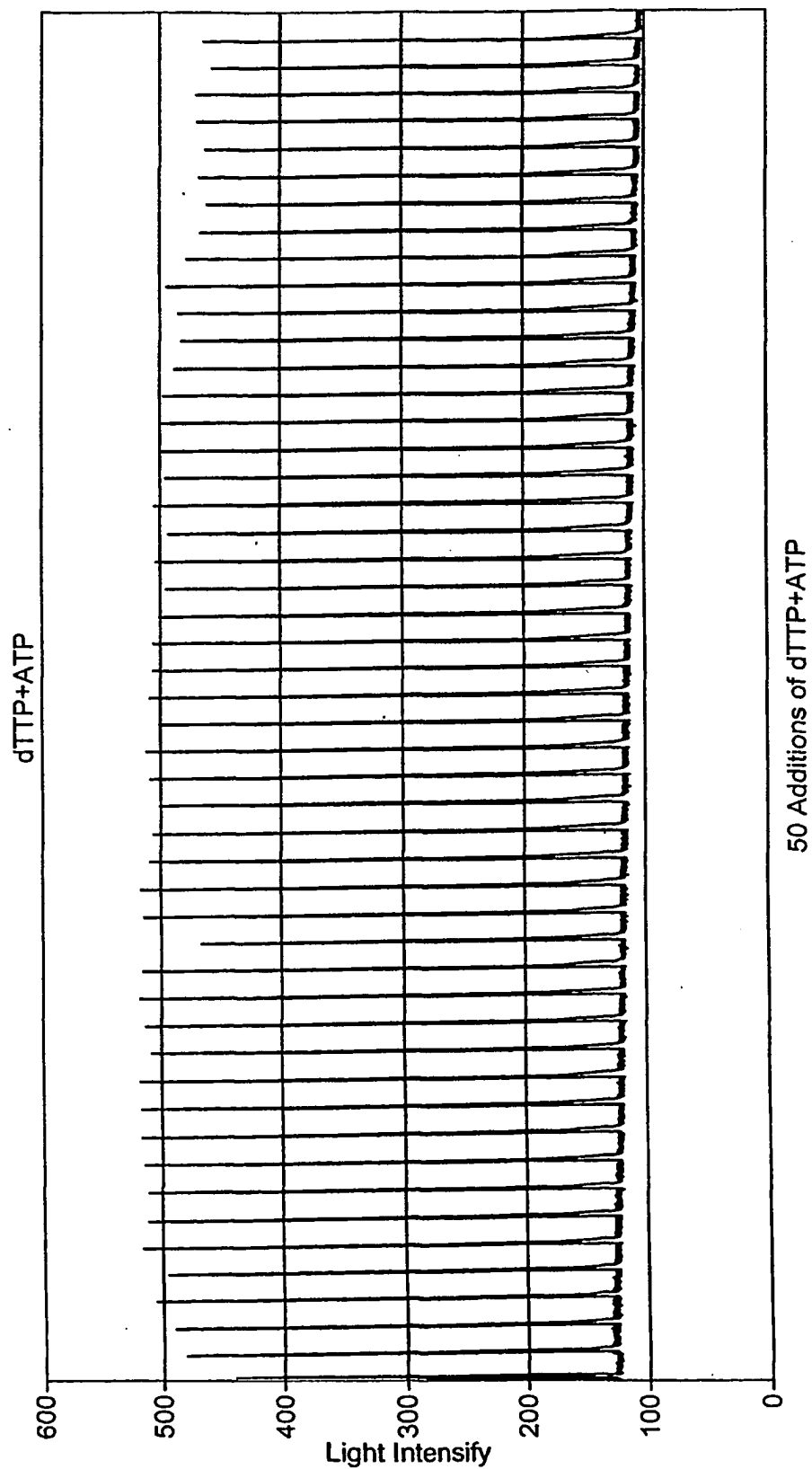


FIG. 4

5 / 22



50 Additions of dTTP+ATP

FIG. 5

6 / 22

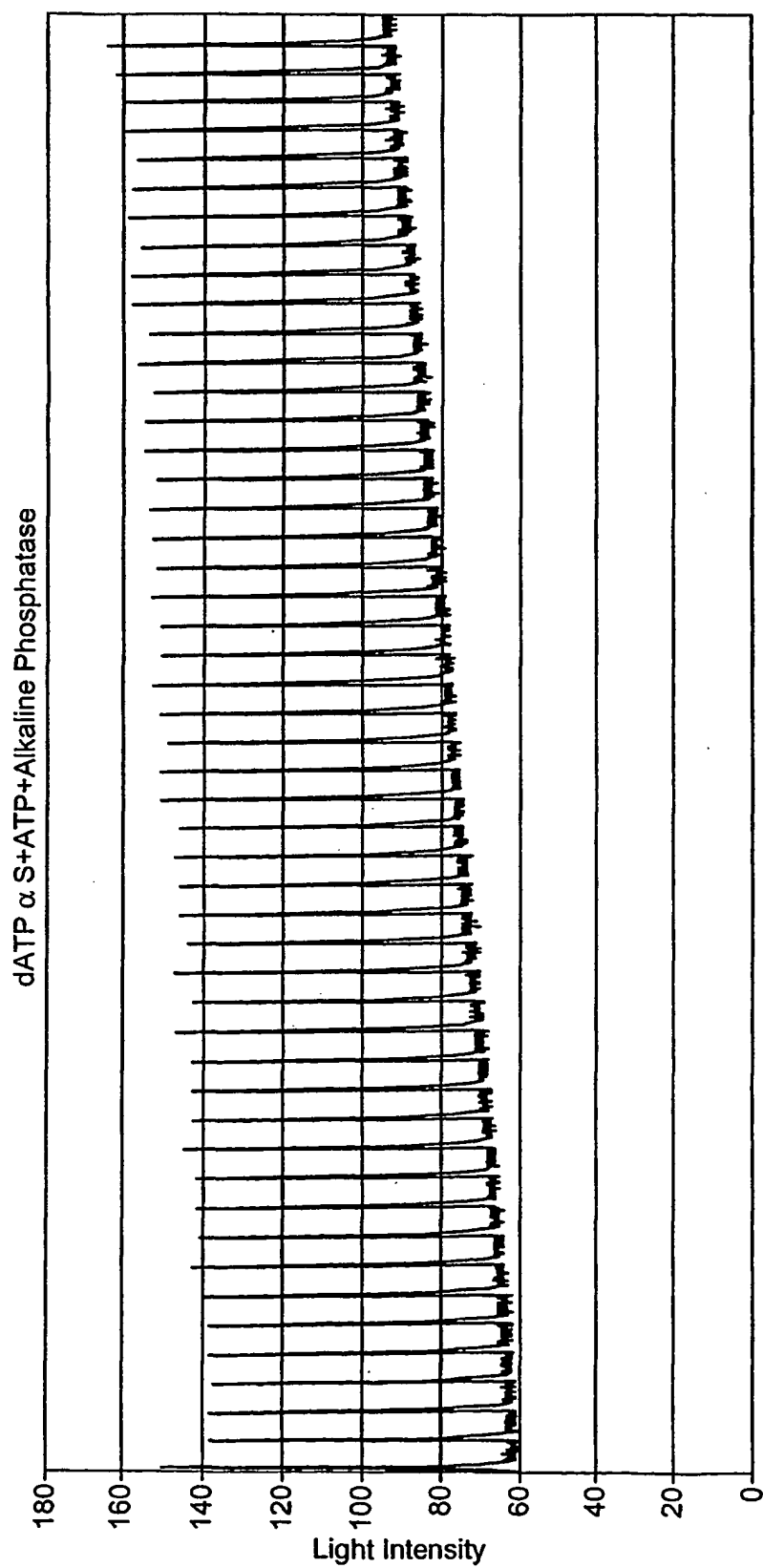


FIG. 6



7 / 22

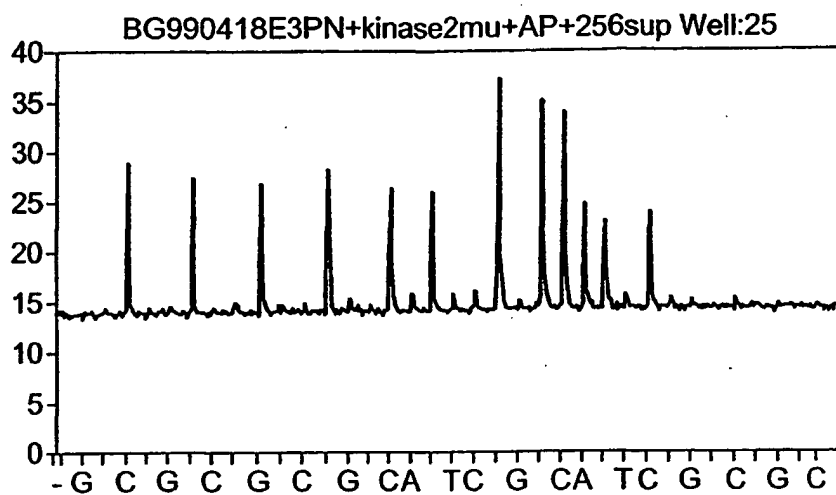


FIG. 7A

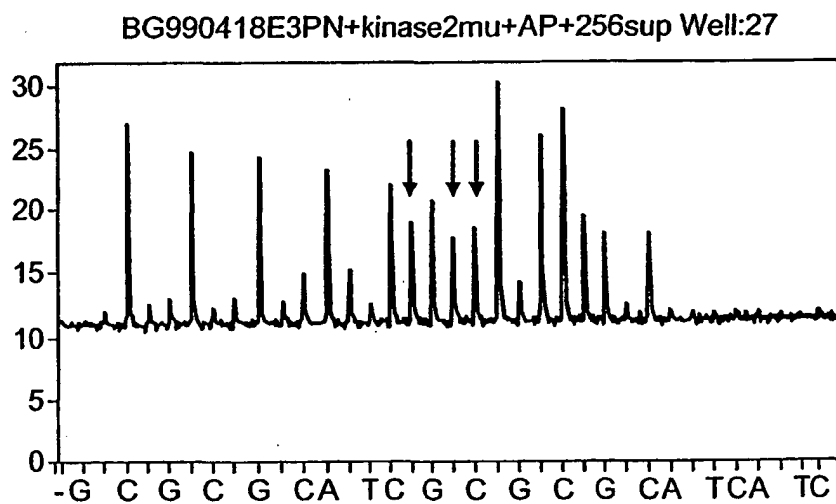


FIG. 7B

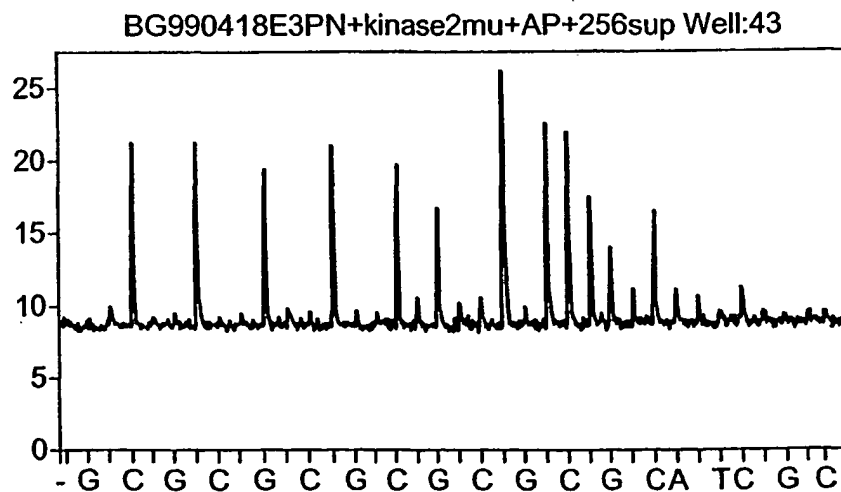


FIG. 7C



9 / 22

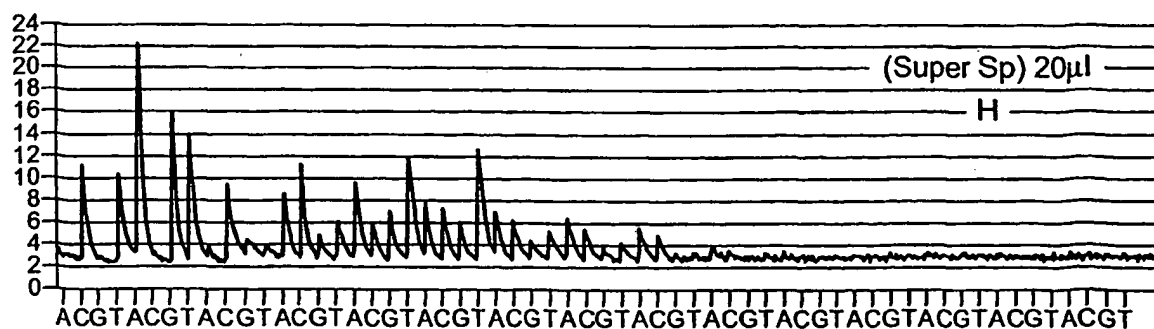
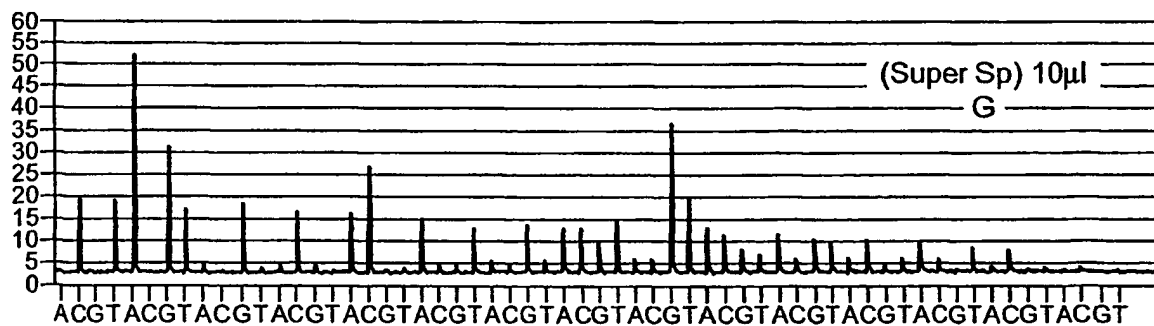
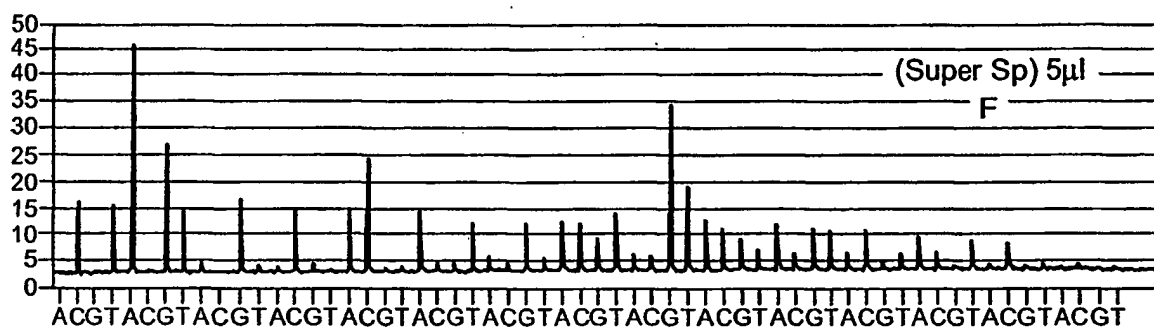
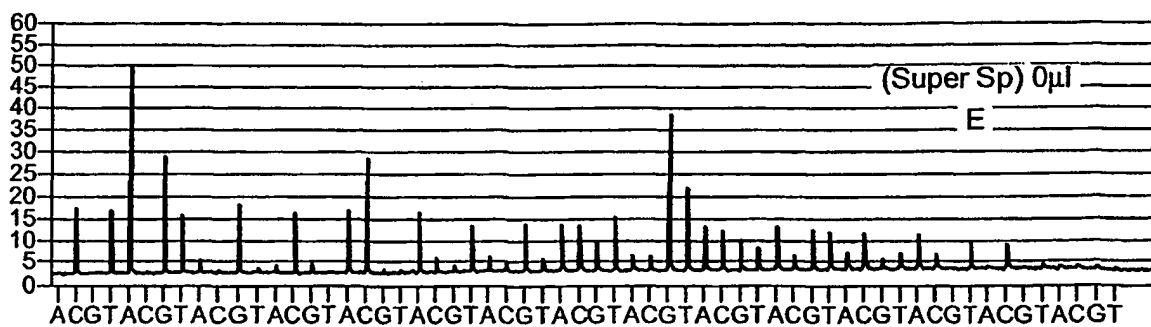
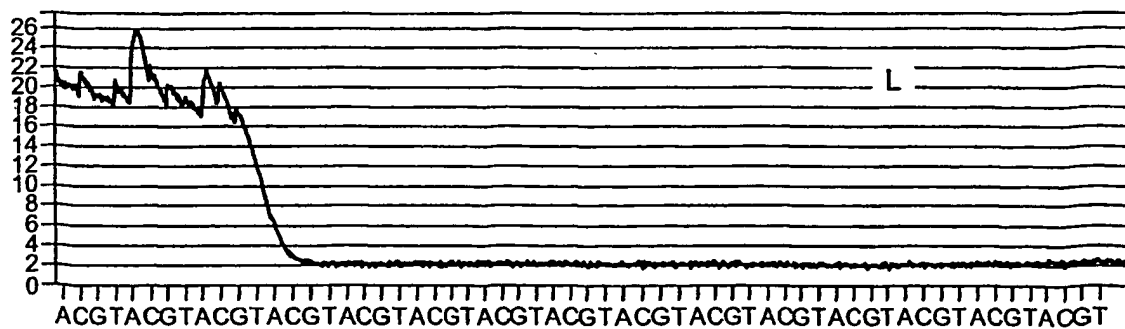
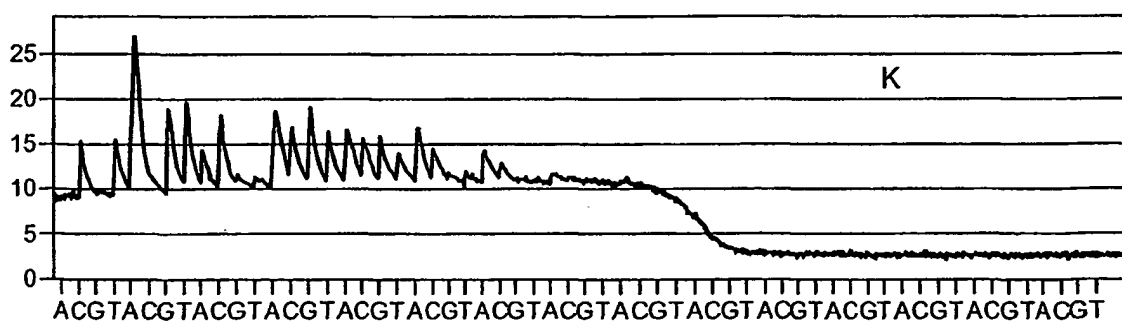
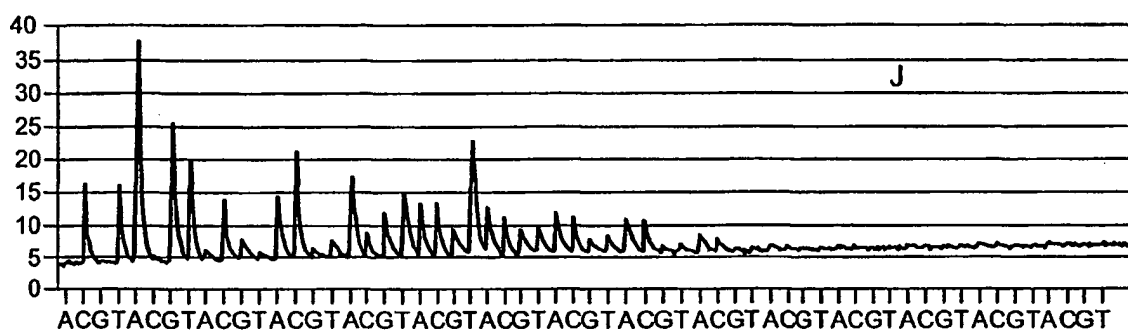
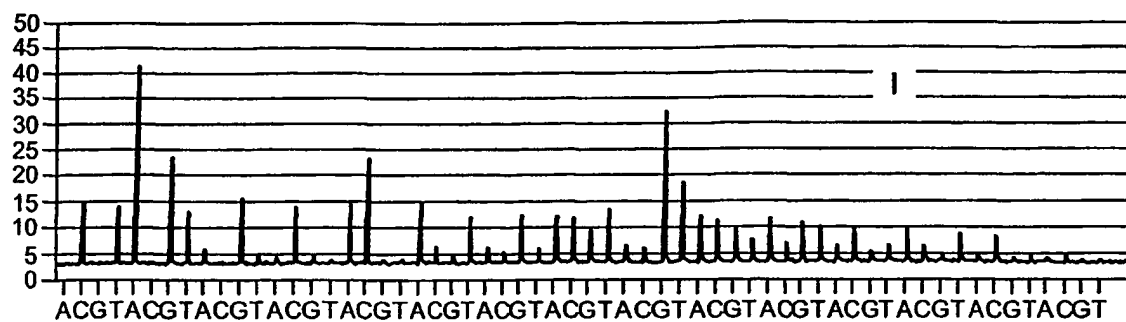


FIG. 8 cont'd

10 / 22



**FIG. 8 cont'd**

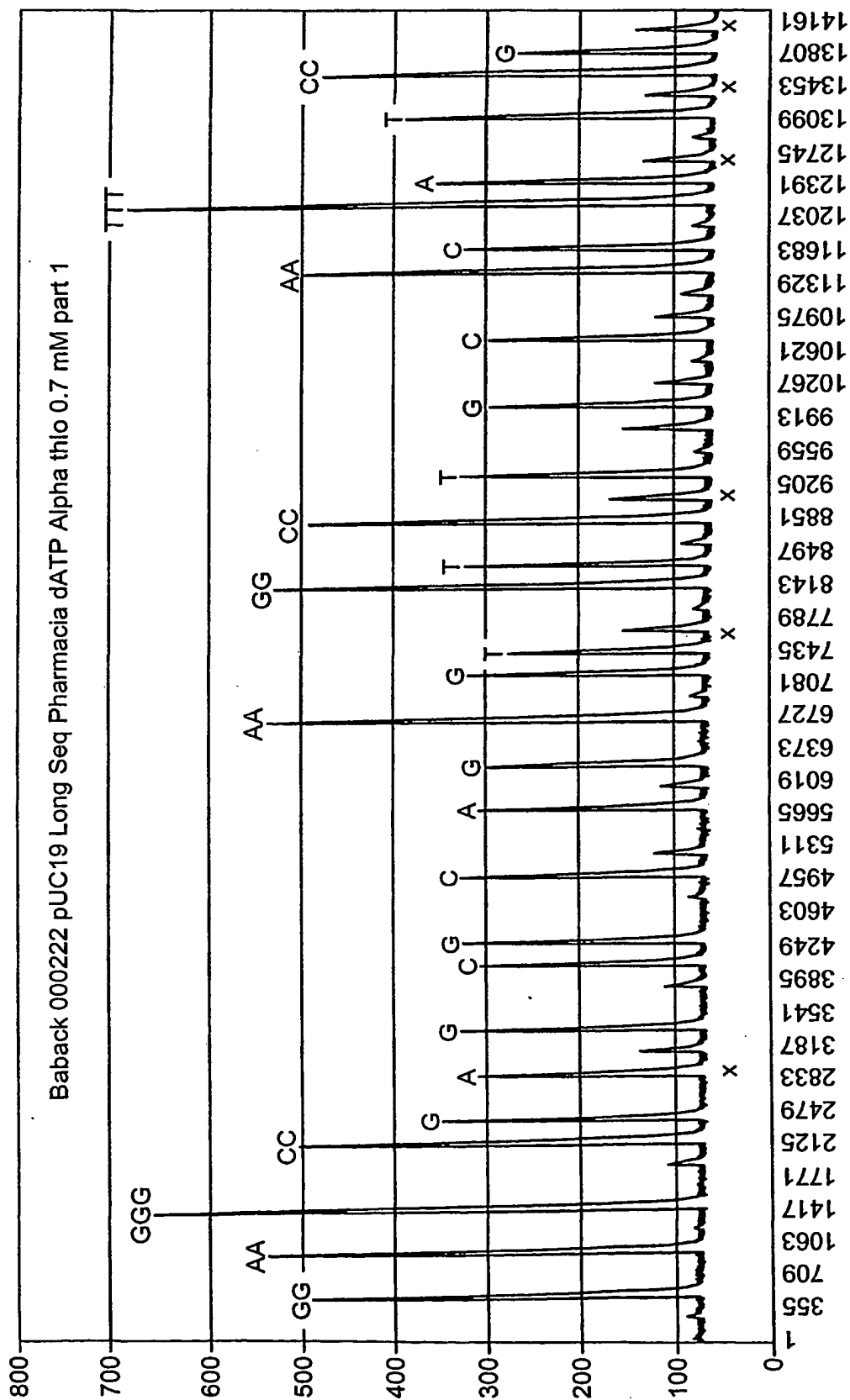


FIG. 9

12 / 22

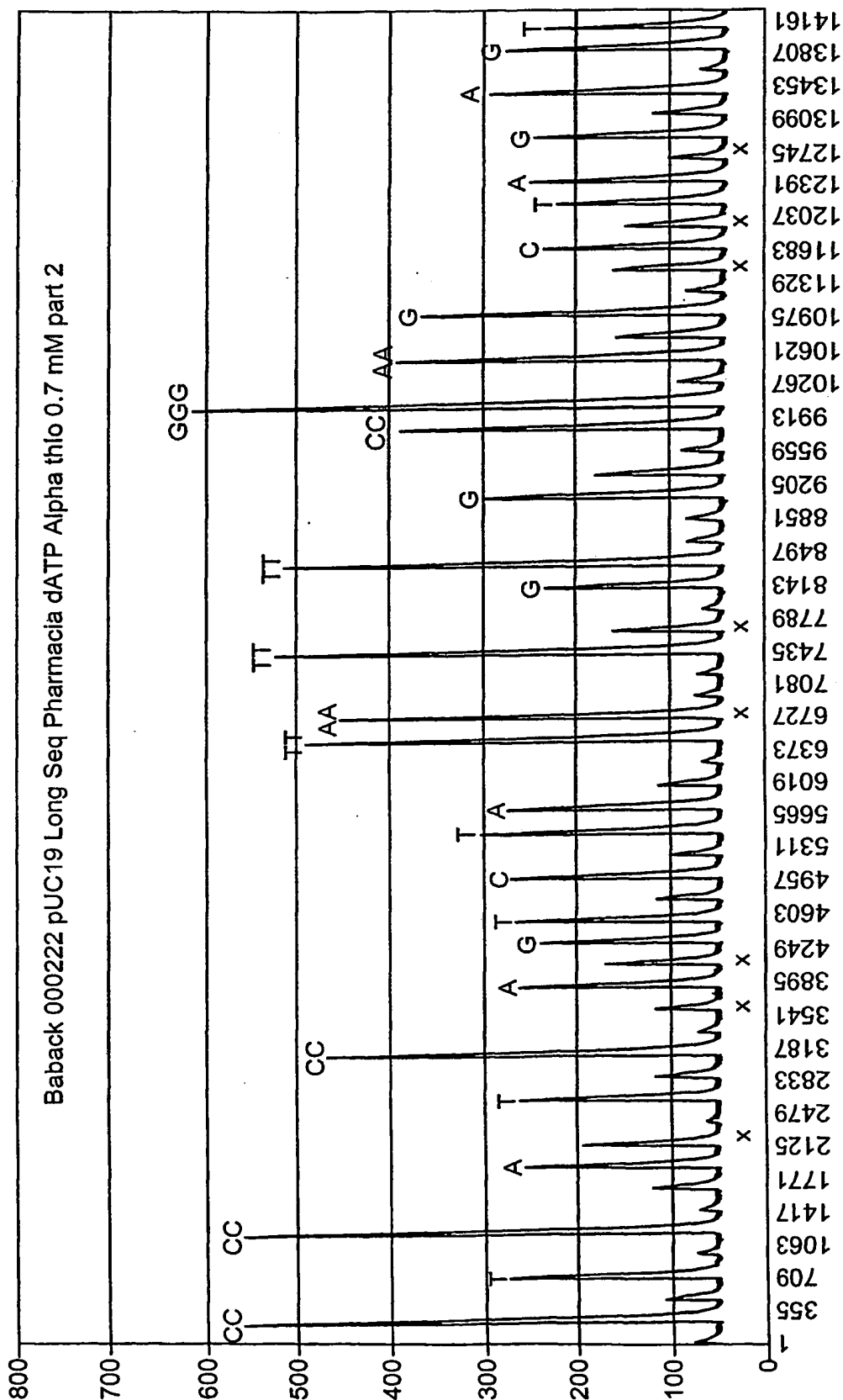


FIG. 9 cont'd

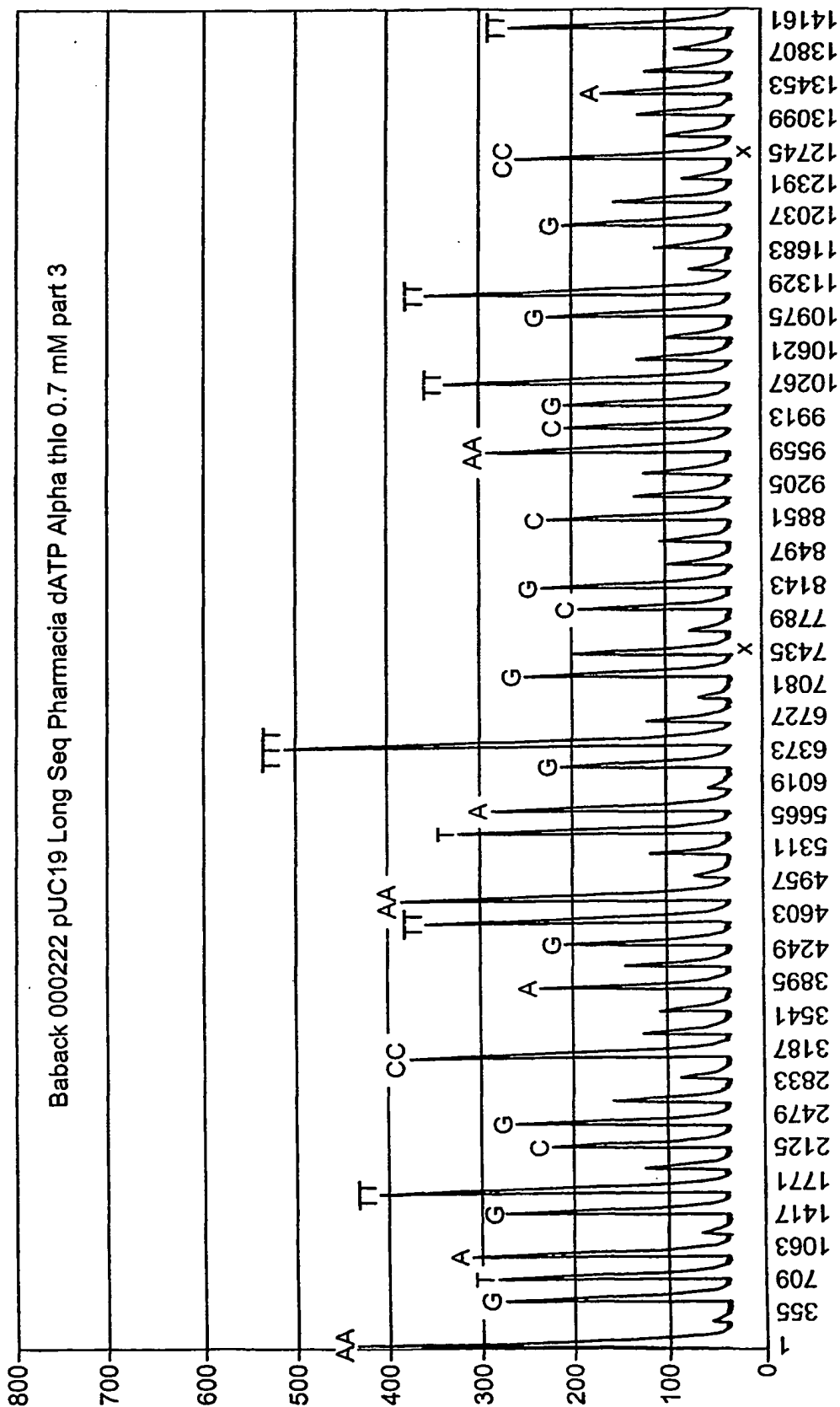


FIG. 9 cont'd

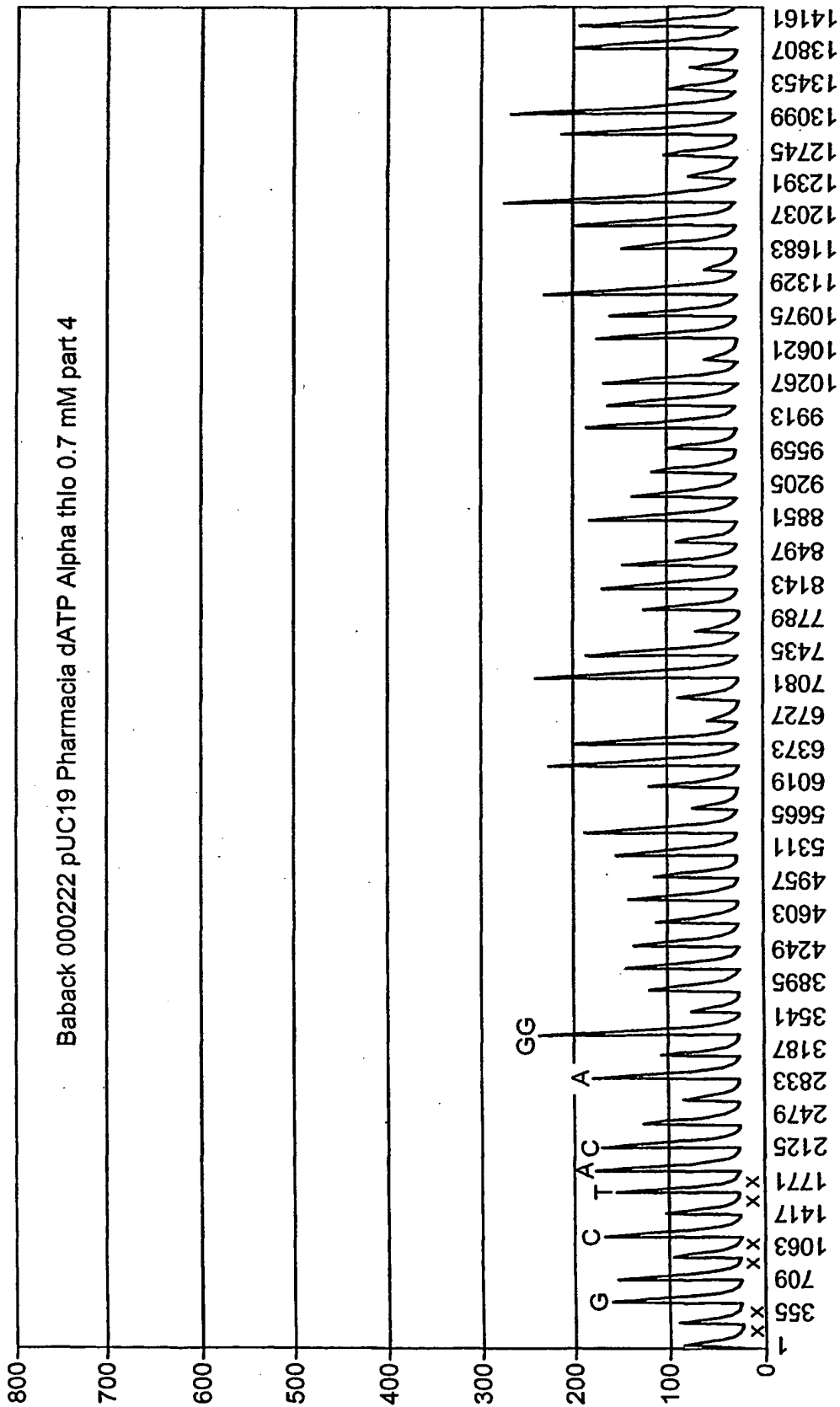


FIG. 9 cont'd



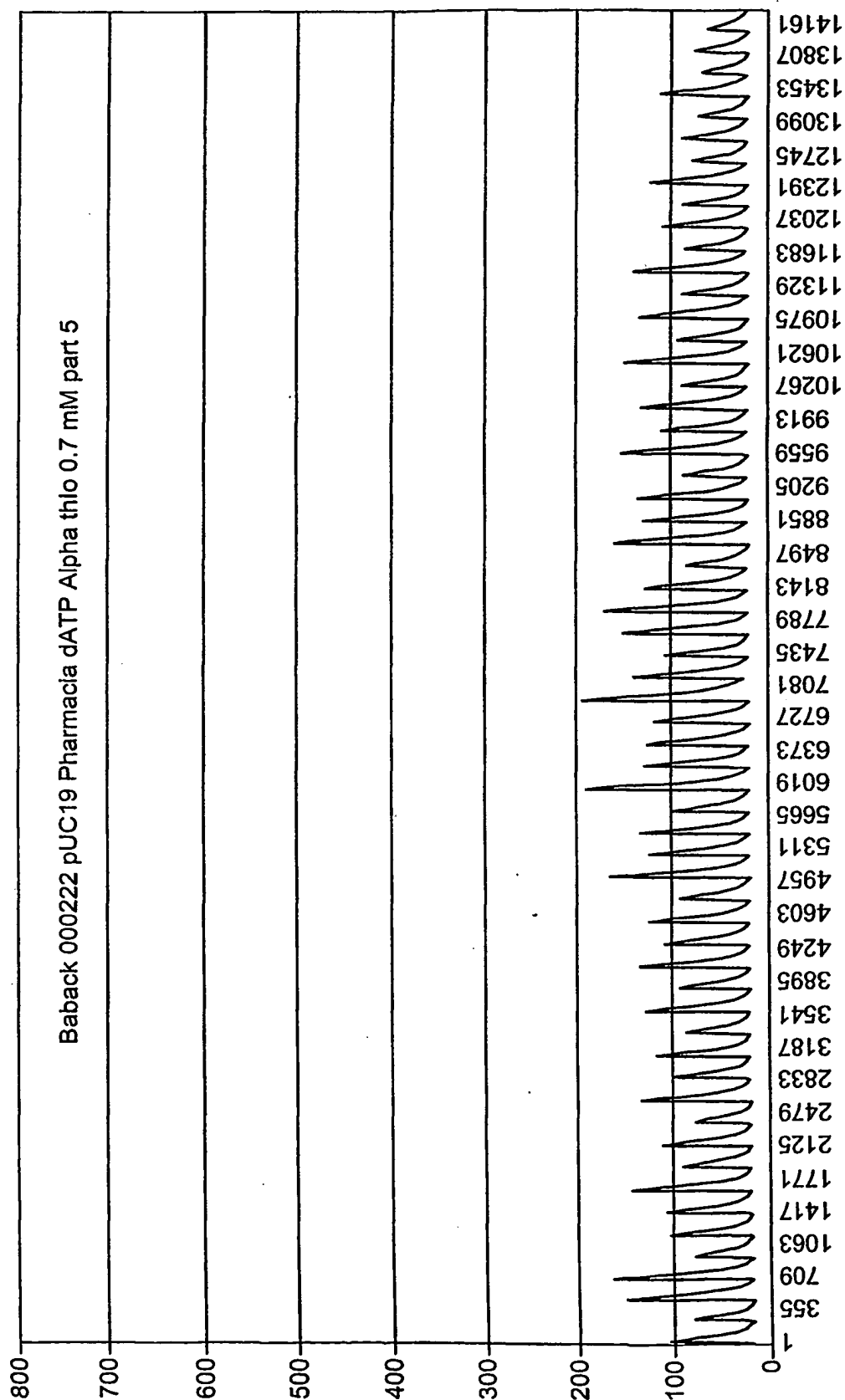


FIG. 9 cont'd

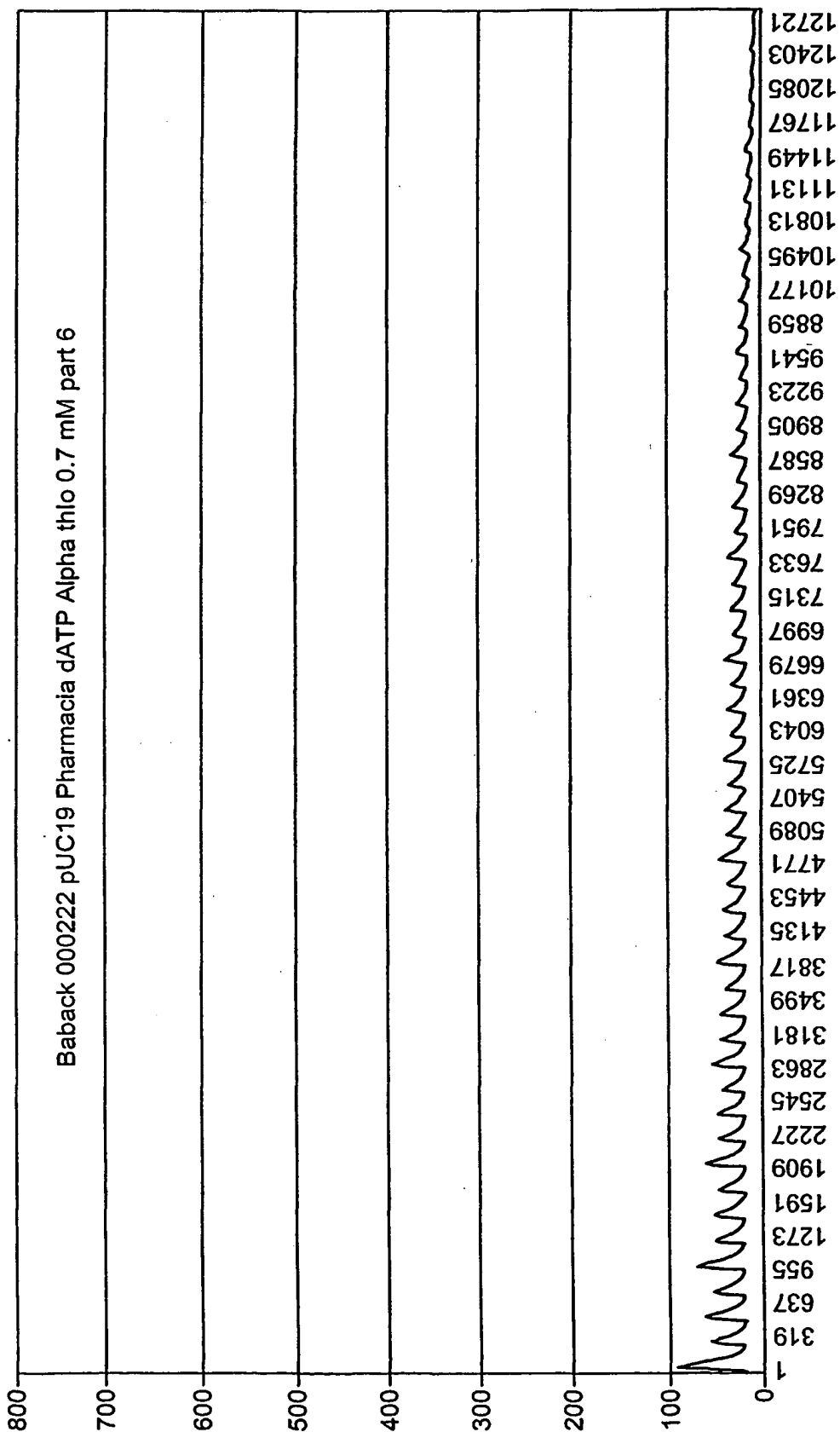
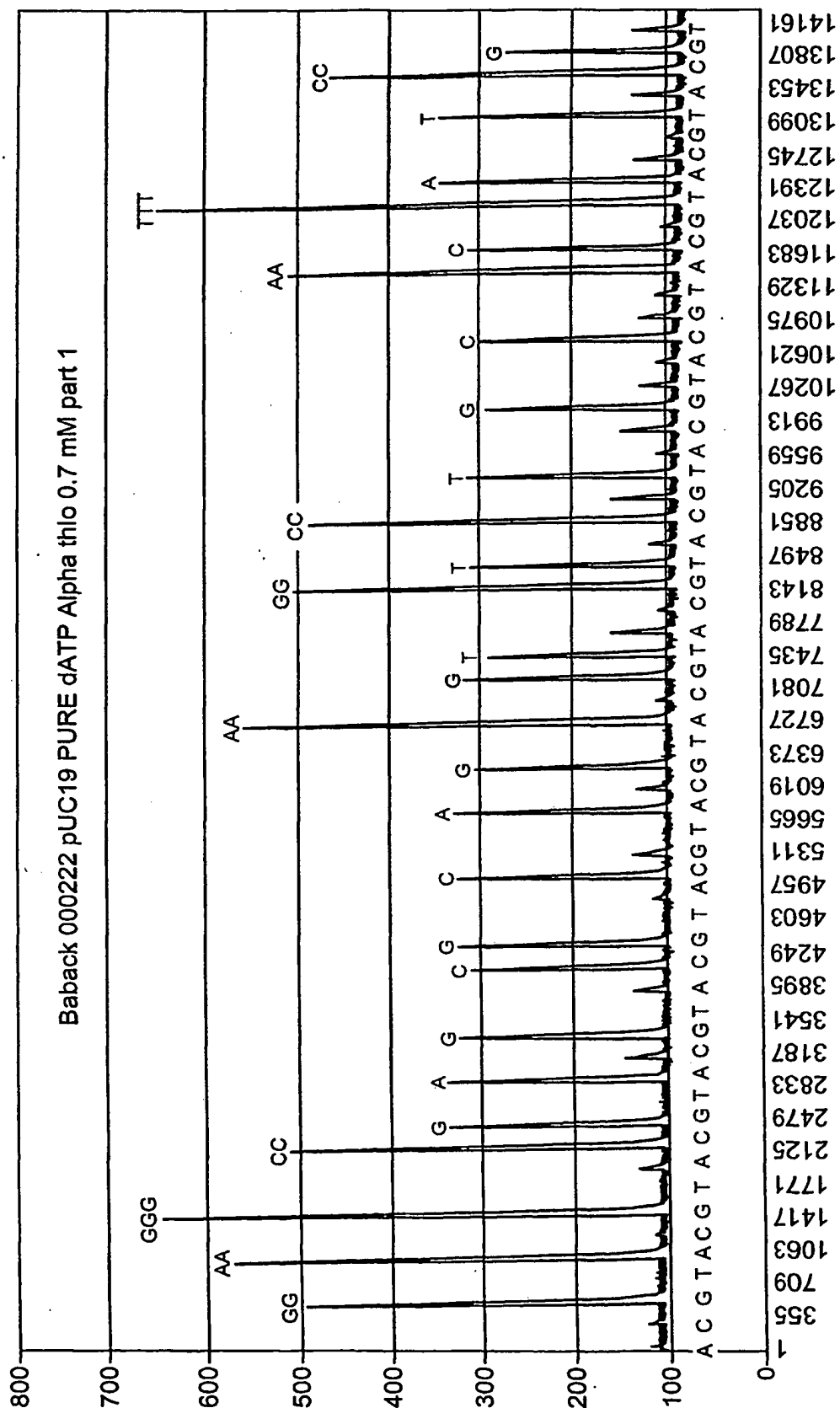


FIG. 9 cont'd



**FIG. 10**

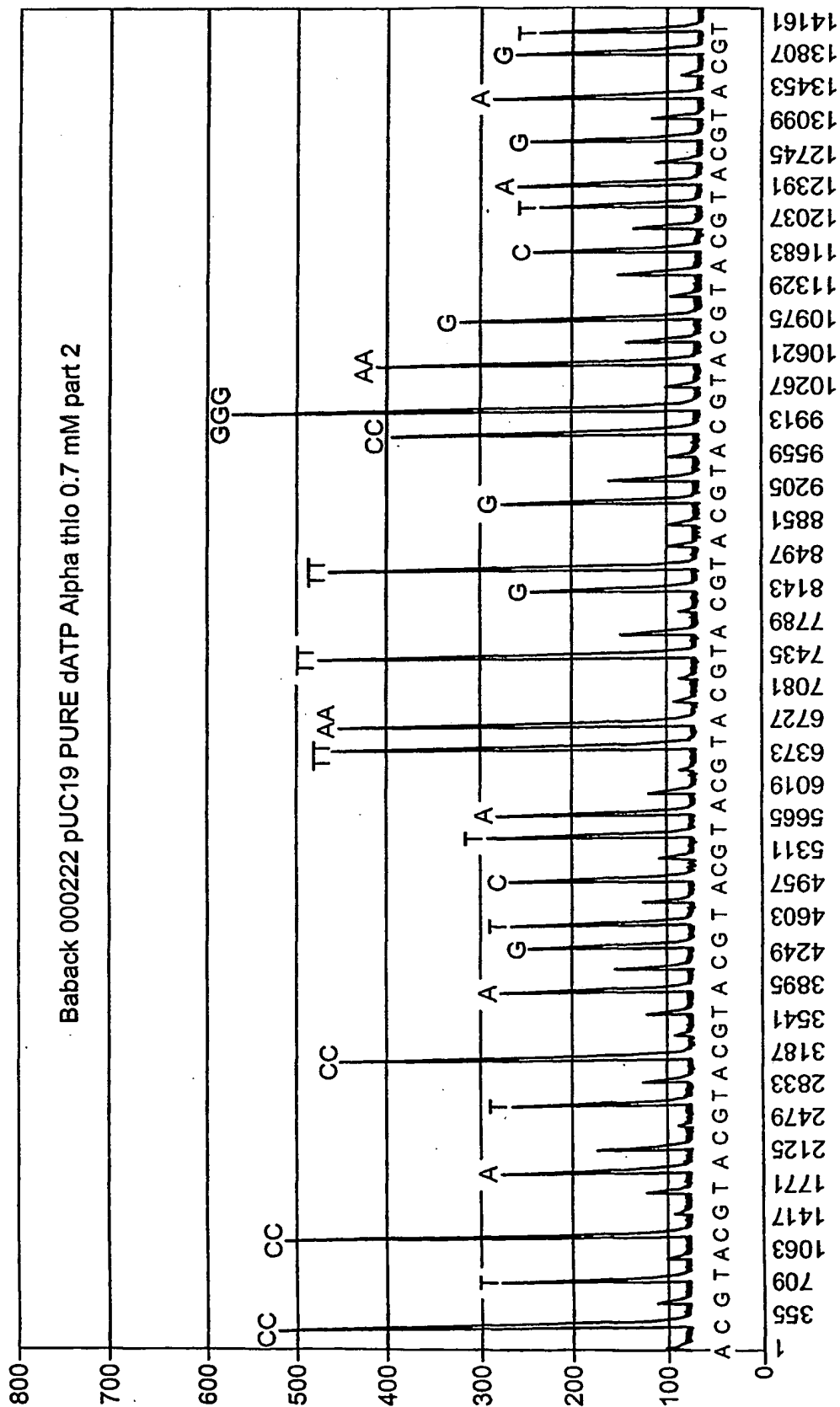
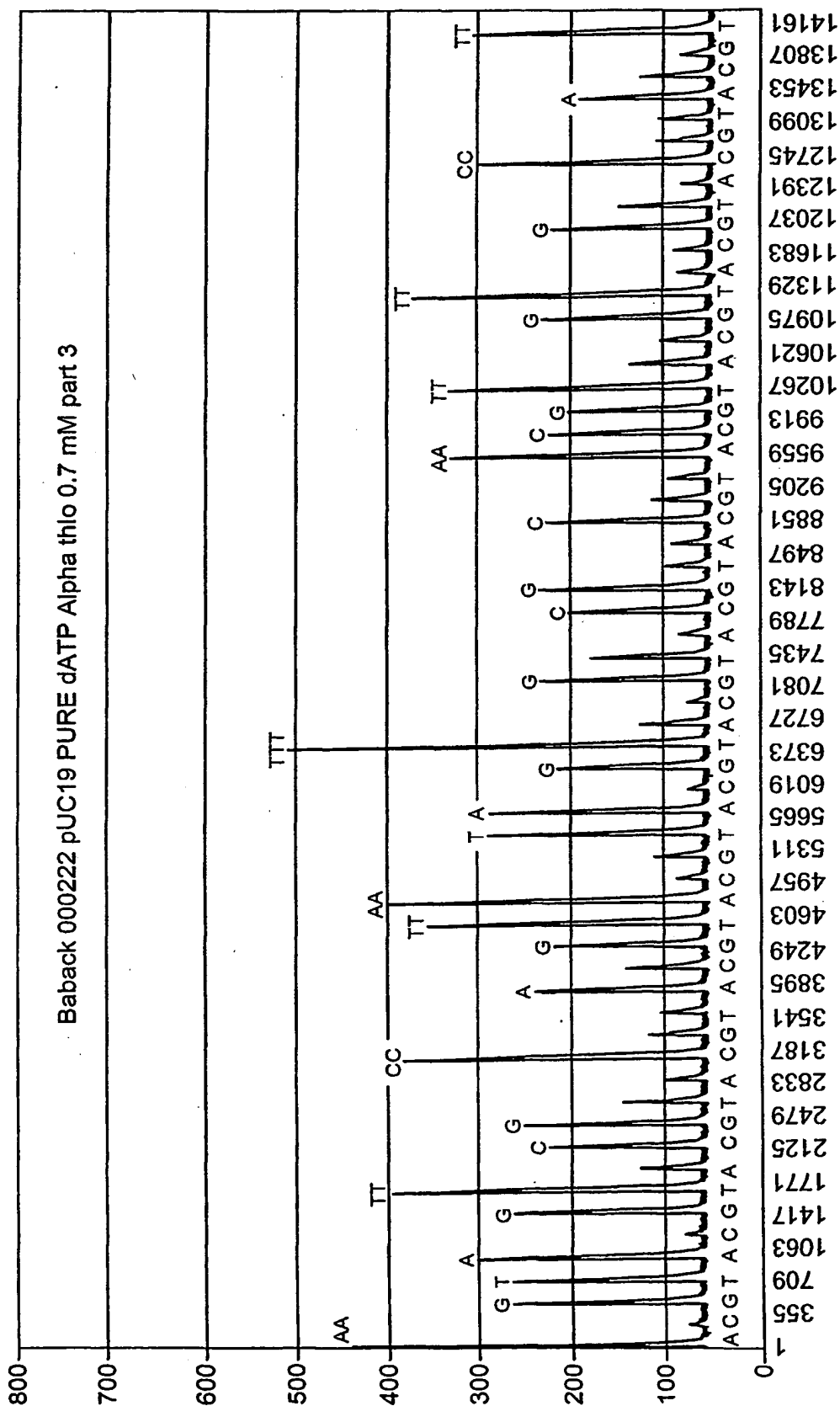
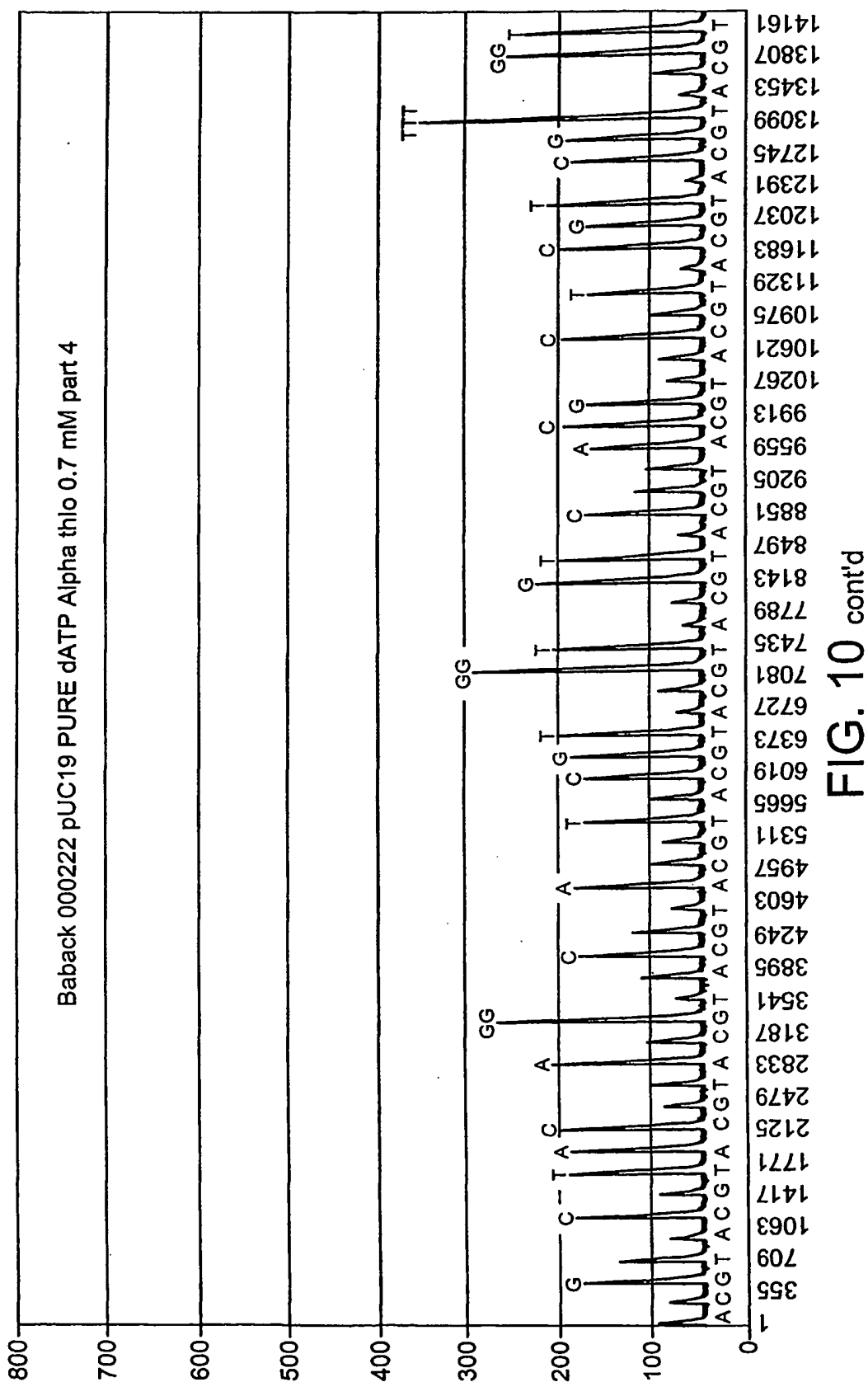
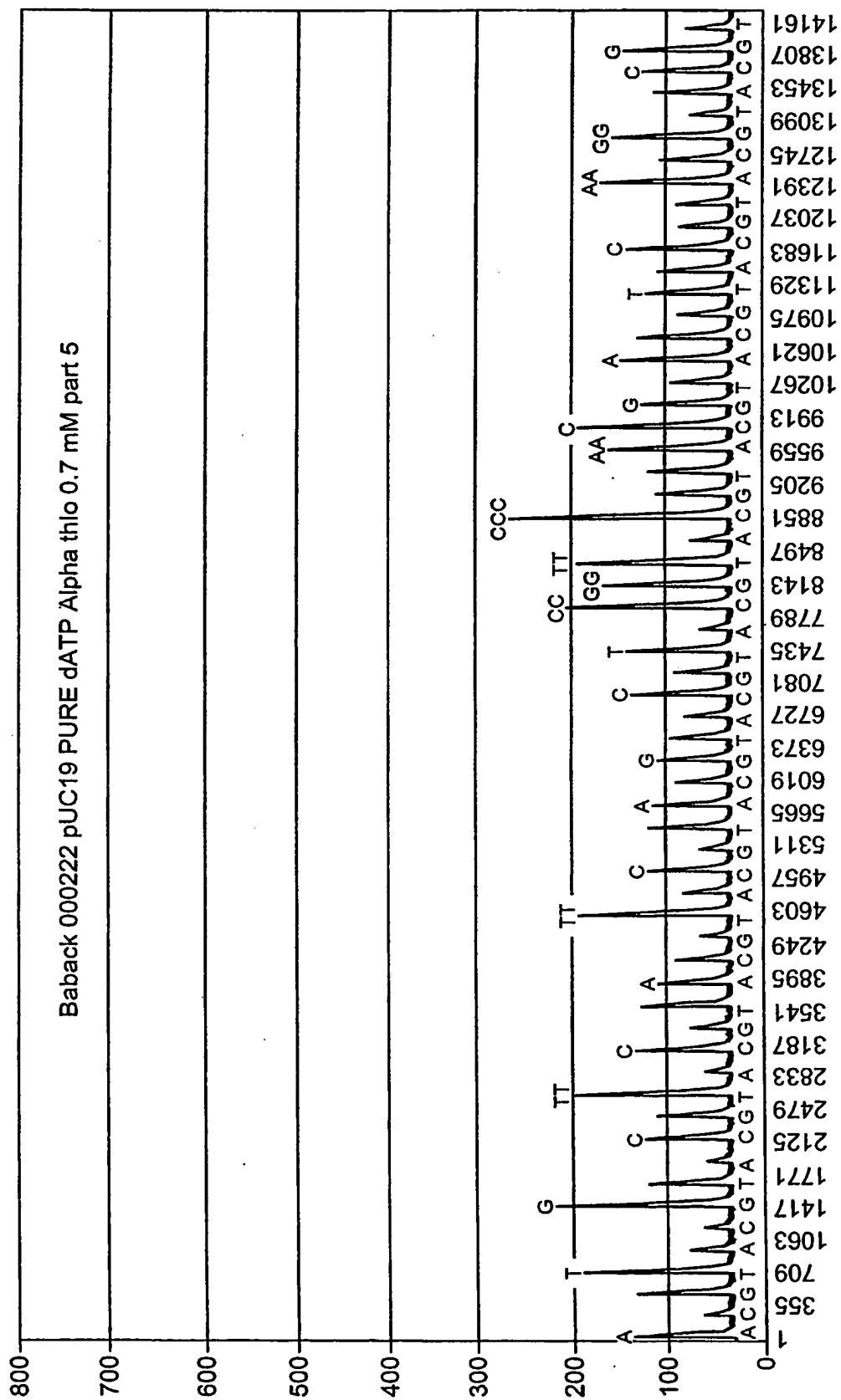


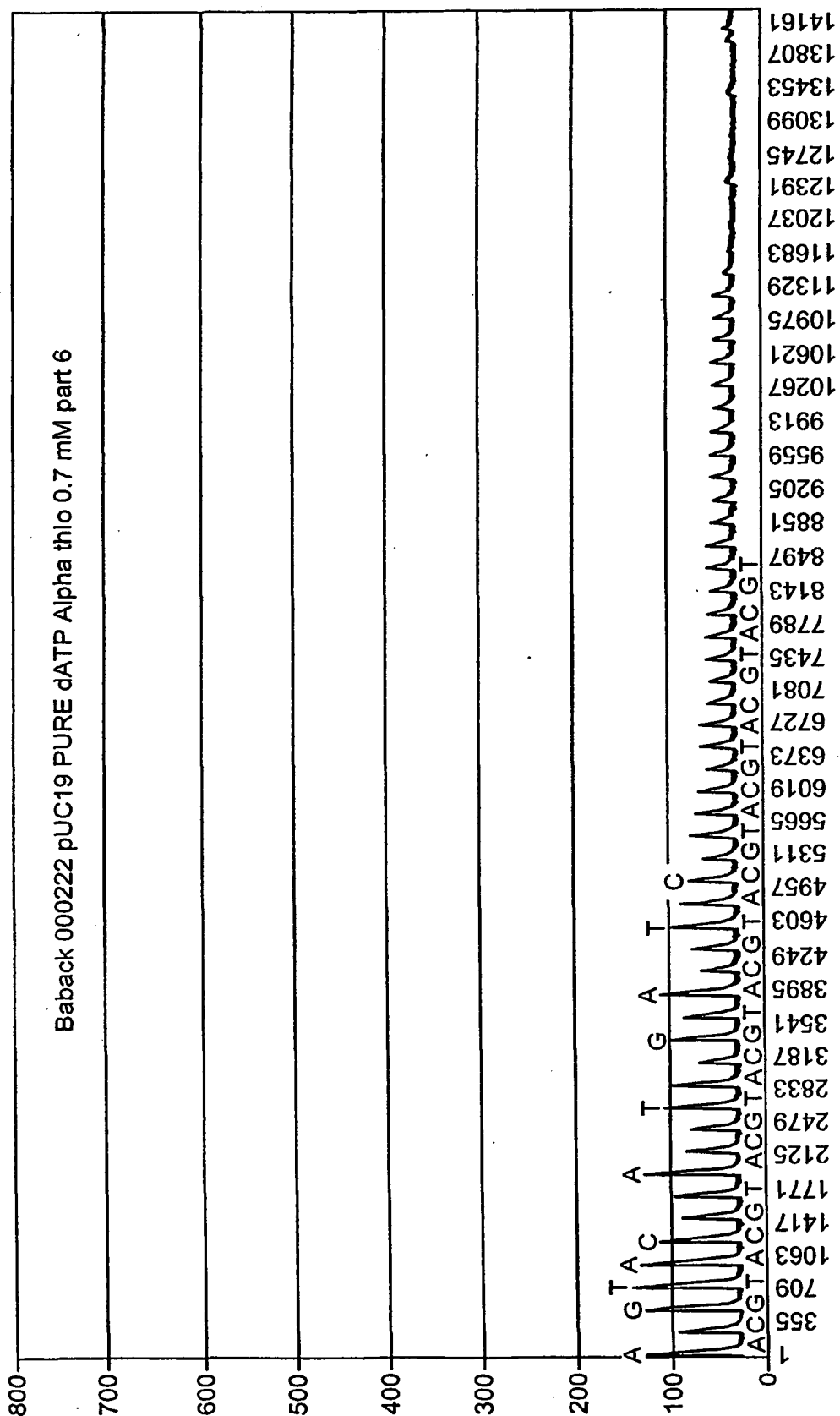
FIG. 10 cont'd







**FIG. 10** cont'd



**FIG. 10** cont'd



(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number  
**WO 02/020836 A3**

- (51) International Patent Classification<sup>7</sup>: C12Q 1/68
- (21) International Application Number: PCT/GB01/04015
- (22) International Filing Date:  
7 September 2001 (07.09.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
0021977.4 7 September 2000 (07.09.2000) GB
- (71) Applicant (for all designated States except US): PY-ROSEQUENCING AB [SE/SE]; Vallongatan 1, S-752 28 Uppsala (SE).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): NYREN, Pål [SE/SE]; c/o Department of Biotechnology, KTH, SCFAB, S-106 91 Stockholm (SE). RONAGHI, Mostafa [SE/US]; Stanford DNA Sequencing & Technology Center, 855 California Avenue, Palo Alto, CA 94304 (US). TALLSJÖ, Annika [SE/SE]; PyroSequencing AB, Vallongatan 1, S-752 28 Uppsala (SE).
- (74) Agents: GARDNER, Rebecca et al.; Frank B. Dehn & Co., 179 Queen Victoria Street, London EC4V 4EL (GB).
- (81) Designated States (national): AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK (utility model), SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— with international search report
- (88) Date of publication of the international search report:  
6 February 2003
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/020836 A3

(54) Title: METHOD OF SEQUENCING DNA

(57) Abstract: The present invention provides a method of identifying a base at a target position in a sample nucleic acid sequence, said method comprising: subjecting a primer hybridised to said sample nucleic acid immediately adjacent to the target position, to a polymerase primer extension reaction in the presence of a nucleotide, whereby the nucleotide will only become incorporated if it is complementary to the base in the target position, and determining whether or not said nucleotide is incorporated by detecting whether Ppi is released, the identity of the target base being determined from the identity of any nucleotide incorporated, wherein, where said nucleotide comprises an adenine base, an  $\alpha$ -thio triphosphate analogue of said nucleotide is used, and the Rp isomer of said analogue and/or the degradation products of said analogue are eliminated from the polymerase reaction step.

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 01/04015

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, CHEM ABS Data, MEDLINE, EMBASE, SCISEARCH, BIOSIS

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	YEE D ET AL: "MECHANISTIC STUDIES ON DEOXYRIBONUCLEIC ACID DEPENDENT RIBONUCLEIC ACID POLYMERASE FROM ESCHERICHIA COLI USING PHOSPHOROTHIOATE ANALOGUES. 1. INITIATION AND PYROPHOSPHATE EXCHANGE REACTIONS" BIOCHEMISTRY, AMERICAN CHEMICAL SOCIETY. EASTON, PA, US, vol. 18, no. 19, 18 September 1979 (1979-09-18), pages 4116-4120, XP000572711 ISSN: 0006-2960	12,13
Y	page 4118 -page 4119 --- -/--	1-11

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*A\* document member of the same patent family

Date of the actual completion of the international search

28 October 2002

Date of mailing of the international search report

08/11/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3018

Authorized officer

Gabriels, J

## INTERNATIONAL SEARCH REPORT

Inte al Application No  
PCT/GB 01/04015

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ROMANIUK P J ET AL: "A STUDY OF THE MECHANISM OF T-4 DNA POLYMERASE WITH DIA STEREOMERIC PHOSPHOROTHIOATE ANALOGS OF DEOXY ATP" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 257, no. 13, 1982, pages 7684-7688, XP002218436 ISSN: 0021-9258 page 7687; figure 3	1-11
Y	ARMSTRONG V W ET AL: "MECHANISTIC STUDIES ON DNA DEPENDENT RNA POLYMERASE FROM ESCHERICHIA-COLI USING PHOSPHOROTHIOATE ANALOGS PART 2 THE ELONGATION REACTION" BIOCHEMISTRY, vol. 18, no. 19, 1979, pages 4120-4123, XP002218628 ISSN: 0006-2960 the whole document	1-11
Y	WO 00 43540 A (GARDNER REBECA ;PYROSEQUENCING AB (SE); RONAGHI MOSTAFA (US)) 27 July 2000 (2000-07-27) claims 1-15; example 5	1-11

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No  
PCT/GB 01/04015

Patent document cited in search report		Publication date		Patent family member(s)	Publication date
WO 0043540	A	27-07-2000	AU	3064200 A	07-08-2000
			EP	1147217 A1	24-10-2001
			WO	0043540 A1	27-07-2000
<hr/>					